5-30-2014 12:00 AM

Statistical Applications in Wildfire Management and Prediction

Lengyi Han
The University of Western Ontario

Supervisor
W. John Braun
The University of Western Ontario

Graduate Program in Statistics and Actuarial Sciences
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

STATISTICAL APPLICATIONS IN WILDFIRE MANAGEMENT AND PREDICTION

(Thesis format: Integrated-Article)

by

Lengyi Han

Graduate Program
in
Statistics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

ABSTRACT

This thesis develops statistical methods and models and applies them to problems related to forest fires. The unifying goal of the work is to provide a data analytic basis for quantifying the uncertainty surrounding fire ignition and fire growth which builds on existing theory where possible.

The main body of the thesis is comprised of three research papers. The Fire Weather Index (FWI) plays an important role in fire management and is central to the first two papers. In the first instance, the block bootstrap confidence interval method is used to deal nonparametrically with the dependence in the FWI data. Because the actual and nominal confidence levels differ substantially, a double bootstrap is applied and used to calibrate the confidence intervals. The calibration technique focuses on interval length-adjustment instead of level-adjustment.

The second paper systematically develops a sequence of parametric time series models for the FWI, starting from some basic physical observations. The final model developed is a seasonal random effect mixture tailed minification model. Numerical approximations using the model allows for calculation of the survival function for the minimum FWI in one or more consecutive days. Tentative results presented here suggest that fire danger prediction may be more effectively accomplished using information on runs of moderately large FWI values, instead of simply using a single-day cut-off.

The third paper studies the fire growth model, Prometheus and re-analyzes the data underlying the associated rate of spread formulas. The main goal of the paper is to incorporate randomness into the deterministic Prometheus model giving rise to a new simulator called Dionysus. The essential idea is a parametric bootstrap. Burn probability contours can be created quickly by Dionysus. These can help fire managers make suppression resource allocation decisions for fires which are currently burning.

Keywords: Fire Weather Index, Prometheus, Bootstrapping, Time Series

CO-AUTHORSHIP STATEMENT

The material presented in this thesis is based on three joint-authored research papers. Versions of Chapers 2 and 4 have been accepted for publication in *Communications in Statistics — Simulation and Computation* and *Environmetrics* and a version of Chapter 3 will be submitted for publication in the near future. I am the lead author in all three papers, but I am grateful to my supervisor, Dr. W. John Braun, for some of the initial ideas and extensive discussions which led to the results conveyed in all three papers. Dr. David Martell should be acknowledged as well for his contribution of the original question concerning the time series analysis of the Fire Weather Index. This is the primary focus of the two of the papers. Cordy Tymstra should also be recognized for the paper on randomizing Prometheus, since he gave very helpful advice and inspiration.

To my Family.

ACKNOWLEDGEMENTS

The research on the modelling of the Fire Weather Index was initiated when I was on a MITACS-Accelerate internship with Alberta Sustainable Resource Development under the supervisorship of Cordy Tymstra. I am grateful to Douglas Woolford for preparing the map of Ontario. The data used is ©1963–2002, Queen's Printer for Ontario, Canada and was referenced under agreement with the Ontario Ministry of Natural Resources.

The research on randomizing Prometheus was supported by Alberta Sustainable Resource Development, MITACS and the Accelerate Internship Program, and GEOIDE, funded by the Government of Canada's NCE program. I am grateful to the Canadian Forest Service for the use of the historic fire behaviour prediction data set, and the Ontario Ministry of Natural resources for the fuel map of the Muskoka Lakes District. I also acknowledge assistance from and helpful discussions with Barry Rowlingson, Mike Wotton, Bob Mazurik, Cordy Tymstra, Rob Bryce and Neal McLoughlin.

I am thankful to my supervisor, John Braun, for his guidance and encouragement through the various stages of my PhD study. Our many stimulating discussions were inspiring and enlightening. There were even many times when he patiently elaborated on the underlying mathematical background of what I was studying, allowing me to see various connections that I might have missed otherwise. I also appreciate the kindness and support of great professors such as David Bellhouse and Reg Kulperger. I should also mention Ward Warkentin who has encouraged my development in data analysis. Good friends have provided additional encouragement and support, and for this I thank Bin Luo, Kevin Wenkai Ma, Paul Nguyen, Zhiying Sun, Tianwei Wen, Jiang Wu, and Jinkun Ken Xiao.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

This thesis is based on three research papers which develop and apply statistical methology in the context of forest fire science. The tools developed in these papers build on and contribute to wildfire science and are readily usable in wildfire management.

Wildfire science is about the nature, causes and effects of fire. Examples of wildfire science are the connection between lightning and fire ignitions, how wave propagation models can be related to fire growth, and what factors contribute to faster growth, such as fuel type, topography and weather. These phenomena have traditionally been understood using deterministic models for fire danger and fire growth.

Of particular importance in Canadian wildfire science are a set of weather-based indices which guide fire management decisions. These indices derive from empirical observation combined with scientific principles. The Canadian Forest Fire Weather Index System (Stocks et al, 1989) outlines the indices and how they are derived from the underlying weather (temperature, wind speed, relative humidity and precipitation). These indices have been incorporated into a fire behaviour prediction system (Forestry Canada, 1992 and Wotton et al, 2009) which is used by fire management agencies across Canada.

The basic level of the system includes the measures Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC) and Drought Code (DC) that quantify, respectively, the wetness of the top layer, moss layer and the deepest layer beneath the forest floor.

The next level of the system includes the Initial Spread Index and the Buildup Index. The Initial Spread Index (ISI) that is related to fire rate of spread (ROS) is calculated from FFMC and DMC. The Buildup Index (BUI) that is related to depth of fire is calculated from DMC and DC.

Finally, the Fire Weather Index (FWI) that is related to fire intensity, or flame length, is calculated from ISI and BUI. Thus, the Fire Weather Index is a statistic that derives from four weather variables that are highly related to fire ignition and spread.

The FWI and ROS will play important roles in the entire thesis. Analysis of the FWI nonparametrically is undertaken in Chapter 2. In Chapter 3, the FWI will be studied as a collection of time series of observations with the goal being to derive a sound statistical model which could be used for simulation purposes as well as testing and prediction. This is important for wildfire management because it can then be used to assist with the development of daily planning models and strategic long term planning models. Such models are used to make decisions about allocation of suppression resources and whether to evacuate a community and so on.

Some fire management agencies are using deterministic fire growth simulators such as Prometheus (Tymstra et al, 2010) to assist with immediate decision-making when there are active wildfires. The Prometheus simulator predicts the future fire shape and area, given curent and forecast weather conditions, together with information about topography and fuel. This is an increasingly important tool which can assist a fire management agency in determining where to deploy fire crews and where potential trouble spots may be located.

Currently, most of the management decisions are based on a deterministic understanding of the fire models and indices. However, the deterministic models built by wildfire science are incapable of managing the ensuing uncertainty following from the chaotic behaviour of wildfire. It is the role of statistics to quantify the uncertainty, to identify when it is large and possibly unmanageable and also to indicate when it is manageable. In Chapter 4, a stochastic version of Prometheus is introduced which can be used to provide probabilistic assessments of the potential for the spread of an existing fire.

This thesis aims to contribute statistical tools which can be used to better understand wildfire and which can be used to assist fire managers in their decision-making processes. The methods developed and studied are block bootstrap for time series, simulation tools both in time and space, and visualization tools. These data analysis tools provide an important unifying theme to this thesis. Time series modelling, both parametric and nonparametric, plays an important role in Chapters 2 and 3, while residual analysis is fundamental to both Chapters 3 and 4, since the results in both of these chapters rest heavily on parametric assumptions.

BIBLIOGRAPHY

[1] Forestry Canada Fire Danger Group. (1992) Development and structure of the Canadian Forest Fire Behavior Prediction System. For. Can., Ottawa, Ont. Inf. Rep. ST-X-3.

[2] Stocks, B.J., Lawson, B.D., Alexander, M.E., Van Wagner, C.E., McAlpine, R.S., Lynham, T.J., Dube, D.E. (1989) Canadian Forest Fire Danger Rating System: an overview. *The Forestry Chronicle*, **65**, 258–265.

[3] Tymstra, C., Bryce, R.W., Wotton, B.M., Taylor, S.W., Armitage, O.B. (2010) Development and structure of Prometheus: the Canadian Wildland Fire Growth Simulation Model . Nat. Resour. Can., Can. For. Serv., North. For. Cent., Edmonton, AB. Inf. Rep. NOR-X-417.

[4] Wotton, B.M., Alexander, M.E. and Taylor, S.W. (2009) Updates and revisions to the 1992 Canadian Forest Fire Behavior Prediction System. For. Can., Sault Ste. Marie, ON Inf. Rep. GLC-X-10.

Chapter 2

A NONPARAMETRIC BOOTSTRAP FOR THE FIRE WEATHER INDEX

2.1 Introduction

The Fire Weather Index (FWI) is a summary measure of fire danger in the Canadian Fire Behaviour Prediction System (Forestry Canada, 1992). It is a function of recent and current temperature, relative humidity, precipitation and wind speed. It is related to moisture content in the surface, DUFF and deep soil layers in the forest floor, and it gives fire managers an indication of the amount of effort (crews, tankers, etc.) that would be required to suppress a fire under such conditions. Larger values of FWI are associated with more effort being required to extinguish a fire.

Martell (1999) developed a Markov chain model for a discretized version of Ontario FWI using five states characterized by specified ranges. That work expressed a need to develop a time series model for the Canadian fire weather index; among other things, a parsimonious time series model would provide a way of characterizing FWI behaviour in different regions. Fujioka and Tsou (1985) used an autoregressive order 2, or AR(2), model for U.S. Fire Danger Ratings. The definition of the U.S. index is somewhat different from the Canadian FWI, so we would not expect completely similar behaviour. In fact, an AR(1) model appears to give an approximately correct model to non-zero Ontario FWI, though the parameter governing this process may be changing from year to year.

The purpose of the present paper is to study the Fire Weather Index for a set of weather stations in Ontario. The approach we take avoids specifying a model explicitly. Instead, we focus on specific characteristics of the FWI process: the first lag autocorrelation, the process mean and the process variance. These parameters can be easily estimated from the data, and a block bootstrap procedure can be used to obtain the corresponding confidence intervals. Such confidence intervals can have very poor coverage properties, however. Our objective is to adjust the confidence intervals coming from a block bootstrap so that coverage probability is closer to the desired (or nominal) level.

First, we will set up our notation. Let θ denote the parameter of interest, and suppose $X_1, X_2, \ldots, X_n$ is the sample coming from a population having parameter θ and distribution function $F(x)$, so that θ is really a functional of F: $\theta(F)$. Note that x can be a scalar if the original observations are independent, but we will also use x as a vector in the case of dependent time series observations. Let $\widehat{\theta}$ denote the estimate of θ based on the sample, i.e. $\theta(\widehat{F})$, where $\widehat{F}(x)$ denotes the empirical cumulative distribution function.

We will consider bootstrap confidence intervals of the form which are often referred to as "percentile" intervals (e.g. Davison and Hinkley, 1997). To compute these intervals, repeatedly draw resamples $X_1^*, X_2^*, \ldots, X_n^*$ from the empirical distribution function $\widehat{F}$. Estimates from each resample are called $\widehat{\widehat{\theta}}$ or $\theta(\widehat{\widehat{F}})$. Here, $\widehat{\widehat{F}}$ denotes the empirical distribution function of the resample, taken from the original sample, thus acting as an estimate of the original empirical distribution. If we have R resamples, then we will have a set of R $\widehat{\widehat{\theta}}$ estimates.

A $1-\alpha$ confidence interval is given by

$$(\widehat{\theta} + \tau_1, \widehat{\theta} + \tau_2) \tag{2.1}$$

where τ_1 is the $\alpha/2$ sample percentile of the distribution of $\widehat{\widehat{\theta}} - \widehat{\theta}$ and τ_2 is the corresponding $1-\alpha/2$ sample percentile.

It is well known that confidence intervals constructed in this way are asymptotically correct to first order under fairly general conditions and provided the data are independent, and there are several ways of making second order corrections (see, e.g. Hall, 1992). However, in cases where the independence assumption fails, higher order asymptotic methods (which fail to account for the incorrectly specified condition) will not provide a correction. The data we are studying exhibits a definite form of dependence, and a block bootstrap (with non-overlapping blocks) will be used to handle this departure from the standard bootstrap assumptions; however, even with block selection methods such as those proposed by Nordman et al (2007), it is difficult to ascertain whether the block bootstrap is accurate. In fact, we will demonstrate that the block bootstrap can be grossly inaccurate. In such situations, a calibration technique may be an appropriate way to make the correction.

We will now use a very simple example to motivate and demonstrate the two forms of calibration to be considered in this paper.

Suppose a random sample of size n has been taken from a normal population with mean μ and variance σ^2, and a parametric bootstrap procedure is used to construct a percentile $1-\alpha$ confidence interval for μ. That is, R_1 resamples are taken from the normal distribution with mean $\bar{X}$ and variance S^2, calculated from the original sample. The average $\hat{\bar{X}}_i$ of the ith resample must then have a normal distribution with mean $\bar{X}$ and variance S^2/n. The percentile confidence interval for μ is given by

$$\bar{X} \pm \tau_{\alpha/2}$$

where τ_α denotes the α-percentile of the distribution of $\hat{\bar{X}}_i - \bar{X}$. If we were able to simulate from this distribution indefinitely, we would be simulating from a normal distribution with mean 0 and variance S^2/n. Therefore, under the exact bootstrap, τ_α is the α-percentile of this normal distribution, and it can be written as $sz_\alpha/\sqrt{n}$, where z_α is a standard normal percentile. In other words, the percentile confidence interval is given by

$$\bar{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \tag{2.2}$$

and its coverage probability is always less than $(1-\alpha)$, although this fact cannot be detected from the computation of the confidence interval only. This result follows from noting that

$$P(\mu \in \bar{X} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}) = P(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{s/\sqrt{n}} \leq z_{\alpha/2})$$

$$< P(-t_{\alpha/2,n-1} \leq \frac{\bar{X}-\mu}{s/\sqrt{n}} \leq t_{\alpha/2,n-1}) = 1-\alpha$$

This example is similar to one given in the nonparametric case by Hall (1992) to motivate the use of a pivotal statistic. Here we use this example to illustrate the calibration methods we will study later.

The calibration methods we will consider both employ a double bootstrap. In the above example, this means we will take re-resamples (of size n) from normal distributions having mean $\hat{\bar{X}}_i$ and variance S_i^2, for $i = 1, 2, \ldots, R_1$. For each of these re-resamples, the average is computed: $\hat{\hat{\bar{X}}}_{i,j}$, $j = 1, 2, \ldots, R_2$, where R_2 denotes the number of re-resamples taken from each of the original normal distributions. The R_1

unadjusted percentile confidence intervals for $\hat{\bar{X}}$ can be computed using

$$\hat{\bar{X}}_i \pm \hat{\tau}_{\alpha/2}$$

where $\hat{\tau}_\alpha$ denotes the α-percentile of the distribution of $\hat{\bar{X}}_{i,j} - \hat{\bar{X}}_i$.

Again, this distribution is normal but now with mean 0 and variance S_i^2/n, so that the ith percentile confidence interval for $\bar{X}$ is really given by

$$\hat{\bar{X}}_i \pm z_{\alpha/2}\frac{s_i}{\sqrt{n}}.$$

Because $\bar{X}$ is known, it is possible to determine that the coverage probability of these intervals is less than $1-\alpha$. Furthermore, we can see that the coverage probability can be corrected to $1-\alpha$ by replacing $z_{\alpha/2}$ by $t_{\alpha/2}$, the $\alpha/2$ percentile of the t-distribution on $n-1$ degrees of freedom. Thus, in this case, the double bootstrap distribution of the statistic

$$\hat{\bar{X}}_{i,j} - \hat{\bar{X}}_i$$

is exactly the same as the bootstrap distribution of the statistic

$$\hat{\bar{X}}_i - \bar{X}.$$

Thus, the double bootstrap tells us exactly how to adjust the original confidence interval in order to obtain correct coverage.

There are at least two points of view that can be taken with this kind of adjustment. Loh's (1991) level adjustment method involves changing α to a new value α' so that the coverage probability of the double bootstrap confidence intervals is actually $1-\alpha$. In our example, this amounts to choosing α' so that $z_{\alpha'} = t_\alpha$. By replacing α with α' in the confidence interval at (2.2), the coverage probability is ensured to be $1-\alpha$.

Another approach to adjusting the percentile confidence interval is to determine the constant c which assures that the proportion of the adjusted double bootstrap intervals

$$\hat{\bar{X}}_i \pm (\hat{\tau}_{\alpha/2} + c)$$

containing $\bar{X}$ is $1-\alpha$. The constant c can then be used to adjust the length of the original confidence interval so that its coverage probability is higher:

$$\bar{X} \pm (\widehat{\tau}_{\alpha/2} + c).$$

For this example, $c = s(t_{\alpha/2} - z_{\alpha/2})/\sqrt{n}$.

2.2 Calibration of Block Bootstrap Confidence Intervals

Unadjusted percentile confidence intervals can have poor coverage properties when used with a block bootstrap. Tables 2.1, 2.5, and 2.9 provide results from a simulation study (fully described in the next section) involving samples of size 100 coming from an AR(1) process. The 6th column shows coverage proportions for the unadjusted percentile 95% confidence interval for the first lag autocorrelation, the process mean and the process variance, respectively. A block bootstrap with various block sizes has been employed.

In some cases, the coverage approaches 90%, but in many cases, coverage is very low, especially where the AR parameter ϕ is large in absolute value. These results show sensitivity to block size, and in some cases, there is an indication that correct specification of the block size will result in reasonable coverage. However, this is not true in all cases.

2.2.1 α Level Adjustment Method

The method is based on a double bootstrap where the α level is changed to a new level α' in an attempt to match the actual coverage with the nominal coverage more closely. This method was first proposed by Loh (1991) for the i.i.d. case. Lee and Lai (2009) studied the method for the dependent case. They also gave asymptotic results.

It is clear that $\widehat{\theta}$ is the true parameter underlying the original resamples. We use the fact that this value is known in a double bootstrap in which confidence intervals for $\widehat{\theta}$ are repeatedly computed and their proportion correct can be calculated. Specifically, random resamples are taken repeatedly from each of the R original resamples. For each of these second level resamples, parameter estimates are again obtained: $\widehat{\widehat{\theta}}$.

$1-\alpha'$ confidence intervals for $\widehat{\theta}$ are of the form

$$(\widehat{\widehat{\theta}}+\widehat{\tau}_1, \widehat{\widehat{\theta}}+\widehat{\tau}_2) \tag{2.3}$$

where $\widehat{\tau}_1$ is the $\alpha'/2$ sample percentile of the distribution of $\widehat{\widehat{\widehat{\theta}}}-\widehat{\widehat{\theta}}$ and $\widehat{\tau}_2$ is the corresponding $1-\alpha'/2$ sample percentile.

The value of α' used in (2.3) is chosen so that a proportion $1-\alpha$ of the resulting confidence intervals contain $\widehat{\theta}$. The α' value represents the adjusted nominal α-level which is then to be used in the confidence interval for θ defined at (2.1). The hope is that the actual confidence level will then be closer to $1-\alpha$. The accuracy and precision of the method will still depend critically on the choice of block size. Some recommendations on how to choose the block size are given at the end of Section 3.

2.2.2 Length Adjustment Method

As an alternative to level adjustment, we propose that the length of the confidence intervals be adjusted directly. This length adjustment method is also based on a double bootstrap. The second bootstrap is used to calculate the constants to be added to the original interval endpoints so that the resulting confidence intervals contain the true parameter with probability $1-\alpha$.

Specifically, we take R_1 resamples, with replacement, from the original sample, computing the statistic of interest, $\widehat{\widehat{\theta}}_i$, for $i=1,2,\ldots,R_1$. From the ith resample, R_2 re-resamples are taken, and the statistic, $\widehat{\widehat{\widehat{\theta}}}_{i,j}$ is computed for the jth re-resample. Confidence intervals for $\widehat{\theta}$ are computed:

$$(\widehat{\widehat{\theta}}_i+\widehat{\tau}_1, \widehat{\widehat{\theta}}_i+\widehat{\tau}_2).$$

where $\widehat{\tau}_1$ is the $\alpha/2$ sample percentile of the distribution of $\widehat{\widehat{\widehat{\theta}}}_{i,j}-\widehat{\widehat{\theta}}_i$ and $\widehat{\tau}_2$ is the corresponding $1-\alpha/2$ sample percentile.

We next find $\widehat{s}_1$ and $\widehat{s}_2$ so that

$$(\widehat{\widehat{\theta}}+\widehat{\tau}_1+\widehat{s}_1, \widehat{\widehat{\theta}}+\widehat{\tau}_2+\widehat{s}_2)$$

contains $\widehat{\theta}$ with probability $1-\alpha$. The length adjusted confidence interval for θ is then

$$(\widehat{\theta}+\tau_1+\widehat{s}_1, \widehat{\theta}+\tau_2+\widehat{s}_2).$$

where τ_1 and τ_2 are the $\alpha/2$ and $1-\alpha/2$ percentiles of the distribution of $\widehat{\widehat{\theta}}-\widehat{\theta}$ as in the original bootstrap confidence interval.

Finding $\widehat{s}_1$ and $\widehat{s}_2$ in practice may be time-consuming. We have considered two methods, one which is slow and accurate, and the other which is a quick approximation.

The first method uses an optimization technique to minimize the quantity

$$\left(\widehat{P}\left(\widehat{\widehat{\theta}}+\widehat{\tau}_1+\widehat{s}_1 \leq \widehat{\theta} \leq \widehat{\widehat{\theta}}+\widehat{\tau}_2+\widehat{s}_2\right)-(1-\alpha)\right)^2$$

with respect to $\widehat{s}_1$ and $\widehat{s}_2$. Here, the notation $\widehat{P}$ indicates that the probabilities we compute are conditioned on the original sample. Using the `optim()` function in R (R Core Team, 2009) under the default settings usually gives satisfactory results, but it can be very slow.

The second method employs a simple independence approximation:

$$1-\alpha \doteq \widehat{P}\left(\widehat{\widehat{\theta}}+\widehat{\tau}_1+\widehat{s}_1 \leq \widehat{\theta}\right)\widehat{P}\left(\widehat{\widehat{\theta}}+\widehat{\tau}_2+\widehat{s}_2 \geq \widehat{\theta}\right).$$

so $\widehat{s}_1$ and $\widehat{s}_2$ can be obtained from

$$\sqrt{1-\alpha} \doteq \widehat{P}\left(\widehat{\widehat{\theta}}+\widehat{\tau}_1+\widehat{s}_1 \leq \widehat{\theta}\right)$$

and

$$\sqrt{1-\alpha} \doteq \widehat{P}\left(\widehat{\widehat{\theta}}+\widehat{\tau}_2+\widehat{s}_2 \geq \widehat{\theta}\right).$$

The latter method leads to faster computation, and can either be used to get approximate intervals, or to obtain initial guesses for use in the former, more accurate, method.

The probability of computing a correct confidence interval which is as narrow as possible still depends on the block size. Some advice on block size selection is given

at the end of the next section.

2.3 Simulation Comparison

In order to assess the length-adjustment calibration method, we conducted a simulation study in which coverage proportion and average confidence interval width was compared for the length-adjusted, level-adjusted and unadjusted methods, respectively.

2.3.1 Study Design

Samples of size 100, 200, 400 and 800 were simulated from order 1 autoregressive processes:

$$X_i = \phi X_{i-1} + \varepsilon_i$$

where the process mean is 0, and the innovation variance is 1. It is well known that for such processes, the first lag autocorrelation ρ coincides with the value ϕ, and the process variance is $1/(1-\phi^2)$. For our simulations, we considered ϕ in the set $\{-.9, -.5, .5, .9\}$. For each parameter setting, 500 samples were generated using the `arima.sim()` function in R (R Core Team, 2009).

For each sample, the first lag autocorrelation $\widehat{\rho}$, average $\bar{X}$ and sample variance S^2 were calculated. Block bootstrap resamples were obtained from each sample, using block sizes of 10, 20, 30 and 40. For each block size, 500 resamples were created, and 95% percentile confidence intervals were constructed for the first lag autocorrelation, process mean and variance.

From each of the resamples, 50 double bootstrap resamples were created using a block size b which was exactly one-half of the original block size (Lee and Lai, 2009). Based on these re-resamples, length-adjusted and level-adjusted 95% percentile confidence intervals were also constructed for the three parameters. To compute the length-adjusted confidence intervals, the independence approximation was used, since the more accurate method was too time-consuming for such an extensive set of simulations.

2.3.2 Results

The results from this simulation study are tabulated in Tables 2.1 to 2.12. Each table lists the parameter being estimated, and the block size being used. The actual coverage proportions for the length adjusted, level adjusted and unadjusted 95% confidence intervals in columns 4 through 6.

The widths of each confidence interval were also calculated. Averages of these widths are listed in Tables 2.1 to 2.12 in columns 7 through 9.

2.3.3 Discussion

From column 4 of Table 2.1, we see that length adjustment gives accurate results for almost any block size that was tried, with the exception of $b = 10$ and $\phi = -.9$, where the coverage proportion was only 69%. Occasionally, the other adjustment method was slightly more accurate, but for highly autocorrelated data, level adjusted confidence intervals and unadjusted confidence intervals can be very inaccurate. Unadjusted coverage proportions can be as low as 3% and level adjustment only corrects this to 27% while length adjustment corrects it to 86%.

Column 7 of Table 2.1 shows that this improvement was made without sacrificing precision. The average confidence interval width for the above case actually decreased from 0.24 (unadjusted) to 0.23 (length-adjusted). At the same time, level adjustment led to an average confidence interval width of 0.40. In general, we see that the average confidence interval width for length adjusted intervals is quite similar to the average width of unadjusted confidence intervals. In one case, the average level adjusted confidence interval was smaller ($b = 10$ and $\phi = -.5$), and the coverage proportion was also higher for level adjustment in that case indicating that level adjustment can work well when there is less dependence and when a good block size is used.

Table 2.5 shows that the confidence intervals for the mean are generally more accurate than for the first lag autocorrelation. Again, length adjustment is usually more accurate than level adjustment, which is not really any more accurate than no adjustment at all. The improvement in accuracy of the length-adjusted intervals has come at a small price. The last 3 columns of Table 2.5 indicate that both adjustment methods give wider confidence intervals. Length-adjusted confidence intervals are slightly wider on average than level-adjusted confidence intervals.

Table 2.9 shows that length adjustment is usually more accurate than level adjustment when estimating the process variance. However, coverage proportions are

not always close to the nominal level (e.g. 62% when $b = 40$ and $\phi = -.5$). Table 2.9 also shows that the length adjusted confidence intervals are wider on average than the level-adjusted confidence intervals which are wider than the unadjusted confidence intervals.

For $n = 200$, we begin to see the effects of poor block size choice when estimating ρ. That is, when $\phi = -.9$, $b = 10$ appears to be too small. The calibration methods offer some improvement in accuracy but are still percentileally incorrect. Larger block sizes work reasonably well and indicate that length adjustment is usually more accurate than the other methods, again without compromising precision (as shown in Table 2.2).

As the sample size increases more, we see that both adjustment methods give accurate confidence intervals for the mean and variance. For cases where $|\phi|$ is large, the unadjusted method is accurate too, but not when $|\phi|$ is smaller. Confidence interval widths are similar, with length adjusted intervals being slightly wider than the others.

When estimating the first lag autocorrelation with large samples, we see that if the block size is too small, then coverage is poor even with adjustment. This is in agreement with asymptotic results such as those given by Nordman et al (2007) which indicate that the block size should increase as a function of sample size. Level adjusted confidence interval are substantially wider than length adjusted confidence intervals which are similar in width to the unadjusted confidence intervals.

The overall conclusion from this simulation study is that length-adjusted confidence intervals for the three types of parameters considered here are almost always more accurate than unadjusted confidence intervals, and they are often more accurate than level-adjusted confidence intervals. When the length-adjusted intervals are less accurate than the level-adjusted intervals, the difference in accuracy is rarely large. This overall gain in accuracy comes at a slight loss in precision relative to the unadjusted confidence intervals, and often at much greater precision than the level-adjusted confidence intervals. Thus, the length-adjusted confidence intervals appear to be more useful from a practical point of view.

2.3.4 Practical Block Size Selection

In addition, the simulation study suggests a practical method for choosing the block size in order to block bootstrap a time series of length n. Time series of length n with

similar statistical characteristics (i.e. autocorrelation structure, variance, etc.) to the observed time series should be simulated repeatedly, and length-adjusted bootstrap confidence intervals should be computed for the parameter of interest, using a range of block sizes. The average width and coverage proportions should be computed in each case. The block size giving the coverage proportion nearest to the nominal coverage value and giving the narrowest intervals should be chosen.

As an example, consider the problem of computing a length-adjusted bootstrap 95% confidence interval for the variance of time series of length $n = 400$ which has properties similar to an AR(1) process with an estimated lag 1 autocorrelation of 0.9.

According to the 4th column of Table 2.3, the coverage proportion of the length-adjusted 95% confidence intervals is 0.95 for block sizes 10, 20 and 30, but not for 40. The 7th column of Table 2.3 shows that the average width of the confidence intervals for block sizes 10 and 20 are 0.52, but only 0.51 for block size 30. Therefore, the narrowest accurate length-adjusted confidence interval should be obtained with block size 30.

Similar reasoning, using columns 5 and 8 of Table 2.11 leads to a block size of 20 for a level-adjusted 95% confidence interval for the variance.

For sample sizes not considered in the simulation study, the tables still provide some guidance in selecting the block size. They can be used to infer coverage proportions and average confidence interval widths for sample sizes within the range of sample sizes considered, and thus, appropriate block sizes. For very large sample sizes, additional simulations would be needed.

In cases where it is not possible to simulate statistically similar time series, one must resort to methods such as those described by Nordman et al (2009).

2.4 Application to the Fire Weather Index

We studied FWI data from six weather stations in Ontario: Kenora, Dryden, Red Lake and Sioux Lookout in the northwest region and Timmins and Temagami in the northeast region. The locations of these six weather stations within the Province of Ontario are shown in Figure 2.1.

The data for each station were restricted to the FWI values for the period from April 29 to September 15 for each year from 1963 through 2004. This period was chosen since it corresponds approximately to the summer fire season in Ontario, and outside of this period, the FWI values are usually at or near 0. During this period,

several different truncations were used by the provincial fire agency to define nil FWI; sometimes, it was defined as any day with FWI less than 0.5 and at other times it was defined as any day with FWI less than 1. In order to achieve consistency from year to year, we defined the FWI to be 0 whenever the calculated FWI was less than 1.0. Less than 1% of the data in the resulting data set was missing, so we felt it unnecessary to handle missingness in a special way, choosing instead to ignore all missing values.

In the FWI data sets, there are many consecutive 0 values. These runs of 0s have been removed in order to obtain time series which are closer to linear, since the probability is zero for a linear process to give rise to more than one 0 in a row. Analyzing the runs of 0 values is of interest, but lies outside the scope of the present paper (see Albet-Green et al. (2013) for such an analysis). For a Markov process, censoring of the runs of 0s can be shown to result in another Markov process. Thus, if Martell's (1999) assumptions are correct, the nonzero FWI process should be Markovian, and a first order autoregressive process could be a useful approximate model.

In order to obtain transformed values of FWI which are more closely approximated by a normal distribution, $\frac{1}{5}$ powers were taken. This power transformation was based on the optimal Box-Cox transformation.

Figure 2.2 shows the trace plot of nonzero FWI for the year 2000 at Red Lake, and it shows the autocorrelation function at Red Lake for all years of data. These plots are not atypical of the data for all six weather stations and show why the AR(1) process could be viewed as a reasonable approximation for the nonzero FWI process.

Thus, the first lag autocorrelation ρ could be a useful summary of the FWI at each weather station. Larger values of ρ correspond to longer periods of either high or low FWI. Consequently, longer periods of high FWI have the potential to result in extreme numbers of fires. A question of immediate interest concerns how ρ might vary geographically. We are also interested in whether ρ is varying temporally, possibly taking on different values each year.

We first conducted bootstrap hypothesis tests for temporal homogeneity of the first lag autocorrelation coefficient ρ at each weather station. The null hypothesis for this test is that ρ is constant between years. The alternative hypothesis is that ρ is different for different years.

The test statistic we use is the variance of the first lag autocorrelation estimates computed for each of the $M = 42$ years of data:

$$S_\rho^2 = \frac{1}{M-1}\sum_{i=1}^{M}(\widehat{\rho}_i - \bar{\widehat{\rho}})^2 \tag{2.4}$$

Under the null hypothesis, this statistic should be close to 0, and under the alternative hypothesis, it should be strictly positive. For the given sample, we denote the statistic value as s_ρ^2.

To construct bootstrap resamples under the null hypothesis, we randomly sample blocks of data from the entire data set with replacement. For the results given here, we used a block size of 20. Enough blocks need to be sampled to construct M "years" of resampled data, each year containing enough values to cover the period from April 29 through September 15: 139 days. By resampling in this way, the data in each year are necessarily identically distributed and should have the same first lag autocorrelation.

Estimates of the first lag autocorrelation are then obtained for each year, and the variance can be computed using the formula at (2.4). By repeatedly resampling in the above way (500 times), we can obtain a bootstrap distribution for the statistic S_ρ^2 under the null hypothesis. A p-value for the test is then computed by calculating the proportion of S_ρ^2 values which exceed s_ρ^2.

Results for each weather station are given in Table 2.13. The p-values for the Northwestern Ontario stations tend to be quite small, and they indicate moderate evidence against the null hypothesis. The stations in Northeastern Ontario have p-values which are very large, indicating no evidence against the null hypothesis. Thus, in the northeastern region, there is no reason to believe that the first lag autocorrelations are changing from year to year, while there appears to be some kind of change occurring in the northwestern region.

We can now calculate length-adjusted and unadjusted bootstrap confidence intervals for the average value of ρ:

$$\frac{1}{M}\sum_{i=1}^{M}\rho_i$$

Our procedure is based on the conservative assumption that the ρ values could possibly differ from year to year which we saw is likely true in some areas though not everywhere.

The first bootstrap sample is constructed by first randomly sampling from the set of years $\{1, 2, \ldots, M\}$ with replacement. From within each of the M selected years of data, bootstrap resamples are constructed using a block bootstrap. For the results given here, we used a block size of 20, and unadjusted percentile 95% confidence

confidence intervals for the average of ρ were computed for each weather station, using 100 resamples. The block size of 20 is partially justified by the simulation results of Table 2.1 which are based on roughly the same sample size, and size the autocorrelations in the FWI data are in the vicinity of 0.5.

To compute the length adjusted confidence intervals, a second bootstrap sample is needed. For each bootstrap resample, the M resampled years are first re-resampled. For each selected year of data, a block bootstrap re-resample is taken using a block size which is one-half the original size, according to the procedure recommended by Lee and Lai (2009). To compute the constants for the length adjusted confidence intervals, we used the more accurate method described in Section 2.2.

The 95% confidence intervals for the average of ρ are given in Table 2.14 for each weather station. It can be seen that the length adjusted intervals are quite different from the unadjusted intervals. Regional differences are not completely apparent here. The Temagami station appears to have somewhat different behaviour from the other stations, but the Timmins station is quite similar to the Northwestern Ontario stations.

To check on the accuracy of the length adjusted confidence intervals, another simulation study was conducted. In this study, 500 samples were simulated. Each sample consisted of $M = 42$ years worth of data, where each year consisted of 200 simulated observations from AR(1) processes. The autoregressive parameter, for each year, was randomly sampled from a normal distribution with mean 0.5 and standard deviation 0.05.

Unadjusted and length adjusted 95% confidence intervals for the average of ρ were computed for each sample, and compared with the true value (0.5, in this case). It was found that 87.8% of the length adjusted confidence intervals were correct, while only 5.6% of the unadjusted intervals were correct. The mean length of the unadjusted confidence intervals was found to be 0.0554 while the mean length of the adjusted confidence intervals was 0.0633.

We also obtained 95% confidence intervals for the variance of ρ using the formula at (2.4). These are given in Table 2.15 for each weather station. Again, the length adjusted intervals are different from the unadjusted intervals. Noteworthy is the fact that the length adjusted confidence intervals for the northeastern stations both contain 0, while the confidence intervals for the northwestern stations do not. This is in agreement with the result from the hypothesis test described earlier.

To check again on the accuracy of the length adjusted confidence intervals, an additional simulation study was conducted. Again, 500 samples were simulated in

exactly the same way as before. Unadjusted and length adjusted 95% confidence intervals for the variance of ρ were computed for each sample, and compared with the true value (0.05^2, in this case). It was found that 94.4% of the length adjusted confidence intervals were correct, while none of the unadjusted intervals were correct. The mean length of the unadjusted confidence intervals was found to be 0.00756 while the mean length of the adjusted confidence intervals was 0.00697.

2.5 Conclusions

In this paper, we have shown that for time series with dependence resembling an AR(1) process, length-adjusted percentile confidence intervals for the process mean, variance and first lag autocorrelation are usually accurate, and often much more accurate than unadjusted confidence intervals.

We applied this methodology to Fire Weather Index data from Ontario and computed confidence intervals for the mean and variance of the annual first lag autocorrelations. We found that the unadjusted confidence intervals tend to underestimate these quantities and that the length adjusted intervals are more accurate and just as precise. We have found evidence that the autocorrelation coefficients change from year to year in the northwestern region of the province, but there is no evidence for such change in the northeastern region. We also found that the first lag autocorrelation behaviour in the extreme northeast region (i.e. Temagami) is somewhat different from the other parts of the province: lower values appear to occur here, so somewhat less extreme fire behaviour might be expected in this area.

In future work, it will be important to apply this methodology to other regions of the boreal forest of Canada.

Further refinement to the methodology will also be addressed in the future. In particular, it should be noted that we did not explicitly use a block size selector in our bootstrap procedure, though methods have been proposed in the literature (e.g. Nordman et al, 2007). The authors have studied another method based on a double bootstrap procedure which works very well, and which will be the subject of a future paper. Use of this procedure gives a block size near 20, which is the value we used for our data analysis.

BIBLIOGRAPHY

[1] Albert-Green, A., Braun, W. J., Martell, D.L. and Woolford, D. G. (2013). *Visualization Tools for Assessing the Markov Property: Sojourn times in the Ontario Fire Weather Index.* Environmentrics. DOI: 10.1002/env.2237

[2] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press.

[3] Forestry Canada Fire Danger Group. (1992). Development and structure of the Canadian Forest Fire behaviour Prediction System. Forestry Canada, Ottawa, Ontario. Information Report ST-X-3.

[4] Fujioka, F.M. and Tsou, T. (1985). Probability modelling of a fire weather index. In *Proceedings of the 8th Conference on Fire and Forest Meteorology.* (Edited by L.R. Donoghue and R.E. Martin) Society of American Foresters, 239–243.

[5] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer, New York.

[6] Lee, S.M.S. and Lai, P.Y. (2009). Double block bootstrap confidence intervals for dependent data. *Biometrika*, **96**, 427–443.

[7] Loh, W.-Y. (1991). Bootstrap calibration for confidence interval construction and selection. *Statistica Sinica*, **1**, 477-491. CMP 92:02

[8] Martell D.L. (1999). A Markov Chain Model of Day to Day Changes in the Canadian Forest Fire Weather Index. *International Journal of Wildland Fire*, **9**, 265–273.

[9] Nordman, D.J., Lahiri, S.N., and Fridley, B.L. (2007). Optimal block size for variance estimation by a spatial block bootstrap method. *Sankhya: The Indian Journal of Statistics Special Issue on Statistics in Biology and Health Sciences*, **69**, 468–493.

[10] R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

ϕ	ρ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	-0.90	10	0.69	0.58	0.18	0.18	0.33	0.21
-0.90	-0.90	20	0.93	0.74	0.55	0.21	0.31	0.19
-0.90	-0.90	30	0.89	0.72	0.57	0.21	0.28	0.18
-0.90	-0.90	40	0.91	0.75	0.63	0.21	0.24	0.17
-0.50	-0.50	10	0.87	0.93	0.88	0.32	0.43	0.31
-0.50	-0.50	20	0.87	0.90	0.87	0.34	0.36	0.29
-0.50	-0.50	30	0.90	0.87	0.86	0.33	0.33	0.27
-0.50	-0.50	40	0.88	0.86	0.84	0.33	0.29	0.25
0.50	0.50	10	0.92	0.92	0.81	0.34	0.49	0.33
0.50	0.50	20	0.92	0.88	0.80	0.36	0.44	0.31
0.50	0.50	30	0.89	0.85	0.75	0.34	0.38	0.28
0.50	0.50	40	0.90	0.83	0.76	0.34	0.34	0.26
0.90	0.90	10	0.86	0.27	0.03	0.23	0.40	0.24
0.90	0.90	20	0.95	0.49	0.25	0.27	0.39	0.24
0.90	0.90	30	0.91	0.45	0.28	0.25	0.35	0.22
0.90	0.90	40	0.91	0.50	0.32	0.25	0.33	0.21

Table 2.1: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the first lag autocorrelation. Sample size is 100.

ϕ	ρ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	-0.90	10	0.41	0.29	0.01	0.12	0.23	0.15
-0.90	-0.90	20	0.83	0.78	0.52	0.14	0.22	0.14
-0.90	-0.90	30	0.86	0.85	0.64	0.14	0.21	0.13
-0.90	-0.90	40	0.93	0.84	0.67	0.14	0.19	0.12
-0.50	-0.50	10	0.84	0.92	0.83	0.24	0.35	0.23
-0.50	-0.50	20	0.91	0.94	0.91	0.25	0.30	0.22
-0.50	-0.50	30	0.90	0.93	0.89	0.25	0.27	0.22
-0.50	-0.50	40	0.89	0.89	0.85	0.23	0.24	0.20
0.50	0.50	10	0.87	0.94	0.81	0.24	0.36	0.24
0.50	0.50	20	0.91	0.90	0.80	0.25	0.33	0.23
0.50	0.50	30	0.89	0.87	0.81	0.25	0.30	0.22
0.50	0.50	40	0.90	0.88	0.83	0.25	0.27	0.21
0.90	0.90	10	0.52	0.13	0.00	0.13	0.25	0.16
0.90	0.90	20	0.90	0.55	0.25	0.16	0.25	0.15
0.90	0.90	30	0.94	0.64	0.42	0.16	0.24	0.15
0.90	0.90	40	0.93	0.65	0.47	0.16	0.23	0.14

Table 2.2: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the first lag autocorrelation. Sample size is 200.

ϕ	ρ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	-0.90	10	0.16	0.01	0.00	0.08	0.16	0.10
-0.90	-0.90	20	0.70	0.65	0.29	0.10	0.16	0.10
-0.90	-0.90	30	0.83	0.82	0.57	0.10	0.15	0.10
-0.90	-0.90	40	0.91	0.87	0.66	0.10	0.14	0.09
-0.50	-0.50	10	0.75	0.97	0.76	0.17	0.26	0.17
-0.50	-0.50	20	0.91	0.96	0.90	0.18	0.24	0.16
-0.50	-0.50	30	0.92	0.96	0.89	0.18	0.21	0.16
-0.50	-0.50	40	0.93	0.93	0.91	0.18	0.19	0.16
0.50	0.50	10	0.79	0.95	0.70	0.17	0.26	0.17
0.50	0.50	20	0.90	0.95	0.84	0.18	0.25	0.17
0.50	0.50	30	0.94	0.95	0.88	0.18	0.23	0.16
0.50	0.50	40	0.94	0.92	0.85	0.18	0.21	0.16
0.90	0.90	10	0.15	0.01	0.00	0.09	0.17	0.11
0.90	0.90	20	0.73	0.54	0.17	0.10	0.16	0.10
0.90	0.90	30	0.87	0.71	0.39	0.11	0.16	0.10
0.90	0.90	40	0.93	0.69	0.47	0.16	0.23	0.14

Table 2.3: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the first lag autocorrelation. Sample size is 400.

ϕ	ρ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	-0.90	10	0.02	0.00	0.00	0.06	0.11	0.07
-0.90	-0.90	20	0.36	0.43	0.09	0.07	0.11	0.07
-0.90	-0.90	30	0.63	0.71	0.42	0.07	0.10	0.07
-0.90	-0.90	40	0.82	0.83	0.57	0.07	0.10	0.07
-0.50	-0.50	10	0.68	0.90	0.61	0.12	0.19	0.12
-0.50	-0.50	20	0.87	0.97	0.87	0.13	0.18	0.12
-0.50	-0.50	30	0.91	0.97	0.90	0.13	0.17	0.12
-0.50	-0.50	40	0.91	0.94	0.88	0.13	0.15	0.12
0.50	0.50	10	0.68	0.87	0.51	0.12	0.19	0.12
0.50	0.50	20	0.86	0.96	0.81	0.13	0.18	0.12
0.50	0.50	30	0.91	0.97	0.89	0.13	0.17	0.12
0.50	0.50	40	0.93	0.96	0.91	0.13	0.16	0.11
0.90	0.90	10	0.01	0.00	0.00	0.06	0.12	0.07
0.90	0.90	20	0.44	0.27	0.04	0.07	0.11	0.07
0.90	0.90	30	0.69	0.65	0.29	0.07	0.11	0.07
0.90	0.90	40	0.92	0.63	0.43	0.16	0.23	0.14

Table 2.4: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the first lag autocorrelation. Sample size is 800.

ϕ	μ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	0.00	10	0.91	0.89	0.86	0.82	0.77	0.67
-0.90	0.00	20	0.90	0.86	0.85	0.79	0.71	0.64
-0.90	0.00	30	0.88	0.84	0.82	0.76	0.64	0.59
-0.90	0.00	40	0.87	0.81	0.80	0.74	0.58	0.55
-0.50	0.00	10	0.83	0.81	0.71	2.70	2.57	1.97
-0.50	0.00	20	0.79	0.74	0.70	3.02	2.62	2.19
-0.50	0.00	30	0.81	0.74	0.71	3.12	2.53	2.23
-0.50	0.00	40	0.81	0.74	0.71	3.00	2.29	2.11
0.50	0.00	10	0.95	0.95	0.94	0.29	0.28	0.26
0.50	0.00	20	0.93	0.91	0.90	0.28	0.25	0.23
0.50	0.00	30	0.89	0.85	0.85	0.27	0.23	0.22
0.50	0.00	40	0.89	0.84	0.83	0.27	0.21	0.20
0.90	0.00	10	0.89	0.97	0.97	0.20	0.26	0.26
0.90	0.00	20	0.95	0.95	0.95	0.27	0.24	0.23
0.90	0.00	30	0.94	0.91	0.91	0.26	0.22	0.21
0.90	0.00	40	0.95	0.90	0.90	0.25	0.20	0.19

Table 2.5: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the mean. Sample size is 100.

ϕ	μ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	0.00	10	0.94	0.94	0.90	0.58	0.56	0.48
-0.90	0.00	20	0.93	0.91	0.89	0.58	0.55	0.49
-0.90	0.00	30	0.90	0.88	0.87	0.57	0.52	0.47
-0.90	0.00	40	0.89	0.86	0.85	0.56	0.50	0.46
-0.50	0.00	10	0.84	0.82	0.70	2.05	2.04	1.53
-0.50	0.00	20	0.89	0.86	0.78	2.43	2.29	1.84
-0.50	0.00	30	0.87	0.83	0.77	2.57	2.29	1.92
-0.50	0.00	40	0.89	0.85	0.81	2.54	2.24	1.92
0.50	0.00	10	0.96	0.96	0.95	0.21	0.20	0.19
0.50	0.00	20	0.94	0.94	0.93	0.20	0.19	0.18
0.50	0.00	30	0.92	0.89	0.88	0.20	0.18	0.17
0.50	0.00	40	0.91	0.90	0.90	0.19	0.17	0.16
0.90	0.00	10	0.90	0.98	0.98	0.13	0.18	0.18
0.90	0.00	20	0.96	0.96	0.95	0.19	0.18	0.17
0.90	0.00	30	0.95	0.95	0.95	0.17	0.16	0.16
0.90	0.00	40	0.93	0.91	0.90	0.17	0.15	0.15

Table 2.6: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the mean. Sample size is 200.

ϕ	μ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	0.00	10	0.96	0.96	0.93	0.42	0.41	0.35
-0.90	0.00	20	0.95	0.94	0.91	0.43	0.41	0.36
-0.90	0.00	30	0.94	0.93	0.90	0.42	0.40	0.36
-0.90	0.00	40	0.94	0.92	0.91	0.41	0.39	0.35
-0.50	0.00	10	0.89	0.88	0.75	1.53	1.55	1.15
-0.50	0.00	20	0.89	0.87	0.80	1.80	1.77	1.38
-0.50	0.00	30	0.92	0.89	0.83	1.92	1.84	1.49
-0.50	0.00	40	0.92	0.91	0.86	1.96	1.86	1.54
0.50	0.00	10	0.96	0.96	0.95	0.15	0.15	0.14
0.50	0.00	20	0.95	0.94	0.92	0.14	0.14	0.13
0.50	0.00	30	0.96	0.94	0.93	0.15	0.14	0.13
0.50	0.00	40	0.93	0.93	0.92	0.14	0.14	0.13
0.90	0.00	10	0.88	0.97	0.97	0.09	0.13	0.13
0.90	0.00	20	0.98	0.97	0.97	0.13	0.13	0.12
0.90	0.00	30	0.96	0.97	0.96	0.12	0.12	0.11
0.90	0.00	40	0.97	0.97	0.97	0.12	0.11	0.11

Table 2.7: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the mean. Sample size is 400.

ϕ	μ	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	0.00	10	0.96	0.96	0.93	0.30	0.29	0.25
-0.90	0.00	20	0.95	0.94	0.93	0.30	0.29	0.26
-0.90	0.00	30	0.95	0.93	0.92	0.30	0.29	0.26
-0.90	0.00	40	0.95	0.93	0.93	0.30	0.29	0.26
-0.50	0.00	10	0.87	0.87	0.75	1.10	1.12	0.83
-0.50	0.00	20	0.93	0.92	0.85	1.32	1.30	1.02
-0.50	0.00	30	0.96	0.95	0.91	1.42	1.40	1.11
-0.50	0.00	40	0.95	0.95	0.89	1.43	1.39	1.14
0.50	0.00	10	0.96	0.96	0.95	0.10	0.10	0.10
0.50	0.00	20	0.96	0.95	0.94	0.10	0.10	0.09
0.50	0.00	30	0.97	0.96	0.95	0.10	0.10	0.09
0.50	0.00	40	0.95	0.94	0.93	0.10	0.10	0.09
0.90	0.00	10	0.86	0.99	0.99	0.06	0.09	0.09
0.90	0.00	20	0.98	0.98	0.97	0.09	0.09	0.09
0.90	0.00	30	0.96	0.96	0.96	0.08	0.08	0.08
0.90	0.00	40	0.97	0.96	0.96	0.08	0.08	0.08

Table 2.8: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the mean. Sample size is 800.

ϕ	$1/(1-\phi^2)$	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	1.33	10	0.89	0.87	0.84	0.93	0.87	0.78
-0.90	1.33	20	0.86	0.86	0.83	0.88	0.80	0.72
-0.90	1.33	30	0.82	0.81	0.78	0.87	0.75	0.69
-0.90	1.33	40	0.82	0.79	0.76	0.83	0.70	0.64
-0.50	5.26	10	0.71	0.64	0.56	5.26	4.94	4.00
-0.50	5.26	20	0.71	0.58	0.54	5.06	4.76	3.99
-0.50	5.26	30	0.66	0.54	0.50	4.77	4.44	3.75
-0.50	5.26	40	0.62	0.55	0.51	4.79	4.49	3.78
0.50	5.26	10	0.82	0.78	0.74	7.39	6.69	5.48
0.50	5.26	20	0.81	0.76	0.74	7.27	6.18	5.44
0.50	5.26	30	0.81	0.76	0.73	6.99	5.59	5.21
0.50	5.26	40	0.80	0.75	0.73	7.00	5.41	5.10
0.90	1.33	10	0.93	0.92	0.89	0.99	0.91	0.82
0.90	1.33	20	0.89	0.86	0.84	0.93	0.82	0.76
0.90	1.33	30	0.86	0.83	0.82	0.95	0.78	0.74
0.90	1.33	40	0.86	0.80	0.79	0.92	0.72	0.68

Table 2.9: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the variance. Sample size is 100.

ϕ	$1/(1-\phi^2)$	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	1.33	10	0.93	0.91	0.89	0.70	0.67	0.60
-0.90	1.33	20	0.92	0.90	0.89	0.68	0.64	0.58
-0.90	1.33	30	0.90	0.87	0.85	0.68	0.61	0.56
-0.90	1.33	40	0.90	0.88	0.86	0.66	0.60	0.55
-0.50	5.26	10	0.80	0.74	0.66	4.80	4.49	3.67
-0.50	5.26	20	0.83	0.77	0.72	5.03	4.68	3.95
-0.50	5.26	30	0.80	0.73	0.69	4.86	4.37	3.83
-0.50	5.26	40	0.83	0.77	0.72	4.87	4.52	3.88
0.50	5.26	10	0.84	0.81	0.74	5.31	5.05	4.05
0.50	5.26	20	0.88	0.85	0.81	6.04	5.50	4.68
0.50	5.26	30	0.91	0.86	0.83	6.07	5.25	4.68
0.50	5.26	40	0.85	0.80	0.78	5.79	4.99	4.49
0.90	1.33	10	0.96	0.94	0.92	0.72	0.68	0.61
0.90	1.33	20	0.91	0.88	0.87	0.72	0.67	0.61
0.90	1.33	30	0.92	0.89	0.88	0.71	0.63	0.58
0.90	1.33	40	0.88	0.85	0.84	0.68	0.60	0.56

Table 2.10: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the variance. Sample size is 200.

ϕ	$1/(1-\phi^2)$	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	1.33	10	0.95	0.93	0.91	0.51	0.49	0.44
-0.90	1.33	20	0.95	0.93	0.91	0.51	0.49	0.44
-0.90	1.33	30	0.95	0.94	0.93	0.50	0.47	0.43
-0.90	1.33	40	0.94	0.91	0.90	0.50	0.47	0.43
-0.50	5.26	10	0.86	0.84	0.75	3.81	3.73	2.94
-0.50	5.26	20	0.86	0.82	0.76	4.01	3.77	3.20
-0.50	5.26	30	0.89	0.80	0.76	4.07	3.79	3.29
-0.50	5.26	40	0.88	0.82	0.78	3.95	3.69	3.22
0.50	5.26	10	0.92	0.91	0.83	4.12	4.02	3.17
0.50	5.26	20	0.92	0.90	0.86	4.45	4.18	3.52
0.50	5.26	30	0.91	0.89	0.87	4.47	4.15	3.59
0.50	5.26	40	0.92	0.89	0.86	4.53	4.10	3.65
0.90	1.33	10	0.95	0.94	0.92	0.52	0.50	0.44
0.90	1.33	20	0.95	0.94	0.93	0.52	0.49	0.45
0.90	1.33	30	0.95	0.93	0.92	0.51	0.48	0.44
0.90	1.33	40	0.92	0.92	0.90	0.51	0.48	0.43

Table 2.11: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the variance. Sample size is 400.

ϕ	$1/(1-\phi^2)$	Block Size	Coverage Proportion			Interval Width		
			length adj.	level adj.	unadj.	length adj.	level adj.	unadj.
-0.90	1.33	10	0.96	0.94	0.92	0.37	0.36	0.32
-0.90	1.33	20	0.96	0.94	0.93	0.36	0.35	0.32
-0.90	1.33	30	0.96	0.94	0.93	0.36	0.35	0.32
-0.90	1.33	40	0.93	0.91	0.90	0.36	0.35	0.31
-0.50	5.26	10	0.91	0.90	0.83	2.89	2.84	2.25
-0.50	5.26	20	0.93	0.90	0.85	3.08	2.97	2.49
-0.50	5.26	30	0.91	0.88	0.84	3.11	2.93	2.54
-0.50	5.26	40	0.93	0.90	0.87	3.21	3.01	2.66
0.50	5.26	10	0.93	0.92	0.87	3.04	3.02	2.36
0.50	5.26	20	0.94	0.91	0.87	3.23	3.10	2.59
0.50	5.26	30	0.95	0.92	0.90	3.38	3.20	2.74
0.50	5.26	40	0.93	0.91	0.88	3.27	3.09	2.70
0.90	1.33	10	0.96	0.95	0.93	0.37	0.35	0.32
0.90	1.33	20	0.96	0.95	0.93	0.37	0.36	0.32
0.90	1.33	30	0.96	0.95	0.94	0.37	0.35	0.32
0.90	1.33	40	0.94	0.94	0.93	0.37	0.35	0.32

Table 2.12: Actual coverage proportions and average widths of length adjusted, level adjusted and unadjusted 95% confidence intervals of the variance. Sample size is 800.

Region	Weather Station	p-value
NW	Red Lake	.038
NW	Sioux Lookout	.044
NW	Kenora	.02
NW	Dryden	.068
NE	Timmins	.234
NE	Temagami	.754

Table 2.13: Results of bootstrap hypothesis test of temporal homogeneity of first lag autocorrelation.

Region	Station	Unadjusted CI	Length adj. CI
NW	Red Lake	(0.296, 0.381)	(0.379, 0.453)
NW	Sioux Lookout	(0.293, 0.393)	(0.369, 0.476)
NW	Kenora	(0.328, 0.405)	(0.401, 0.465)
NW	Dryden	(0.338, 0.420)	(0.421, 0.500)
NE	Timmins	(0.313, 0.401)	(0.392, 0.478)
NE	Temagami	(0.218, 0.349)	(0.295, 0.420)

Table 2.14: 95% confidence intervals for mean of the first lag autocorrelation parameter at each of 6 weather stations in Northwestern (top 4) and Northeastern Ontario (bottom 2).

Region	Station	Unadjusted CI	Length adj. CI
NW	Red Lake	(0.0145, 0.0334)	(0.00371, 0.0194)
NW	Sioux Lookout	(0.0168, 0.0387)	(0.00128, 0.0203)
NW	Kenora	(0.0134, 0.0295)	(0.00295, 0.01787)
NW	Dryden	(0.0122, 0.0270)	(0.000541, 0.015225)
NE	Timmins	(0.0138, 0.0324)	(-0.000413, 0.014026)
NE	Temagami	(0.00903, 0.04717)	(-0.0230, 0.0187)

Table 2.15: 95% confidence intervals for variance of the first lag autocorrelation parameter at each of 6 weather stations in Northwestern and Northeastern Ontario.

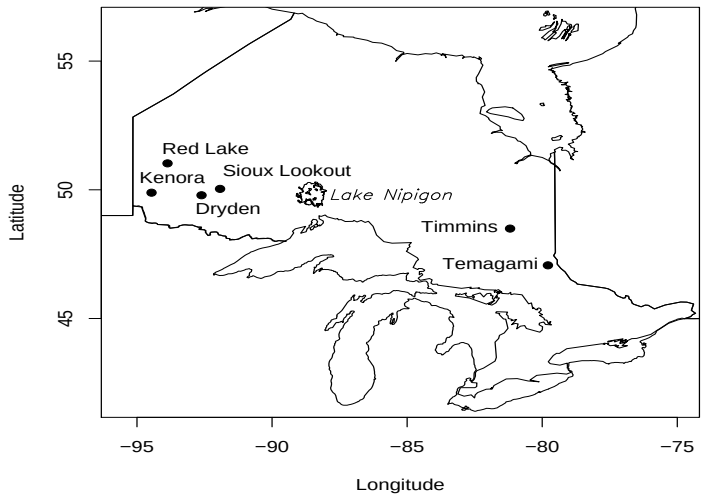

Figure 2.1: Geographical locations of the weather stations: sources of the FWI data.

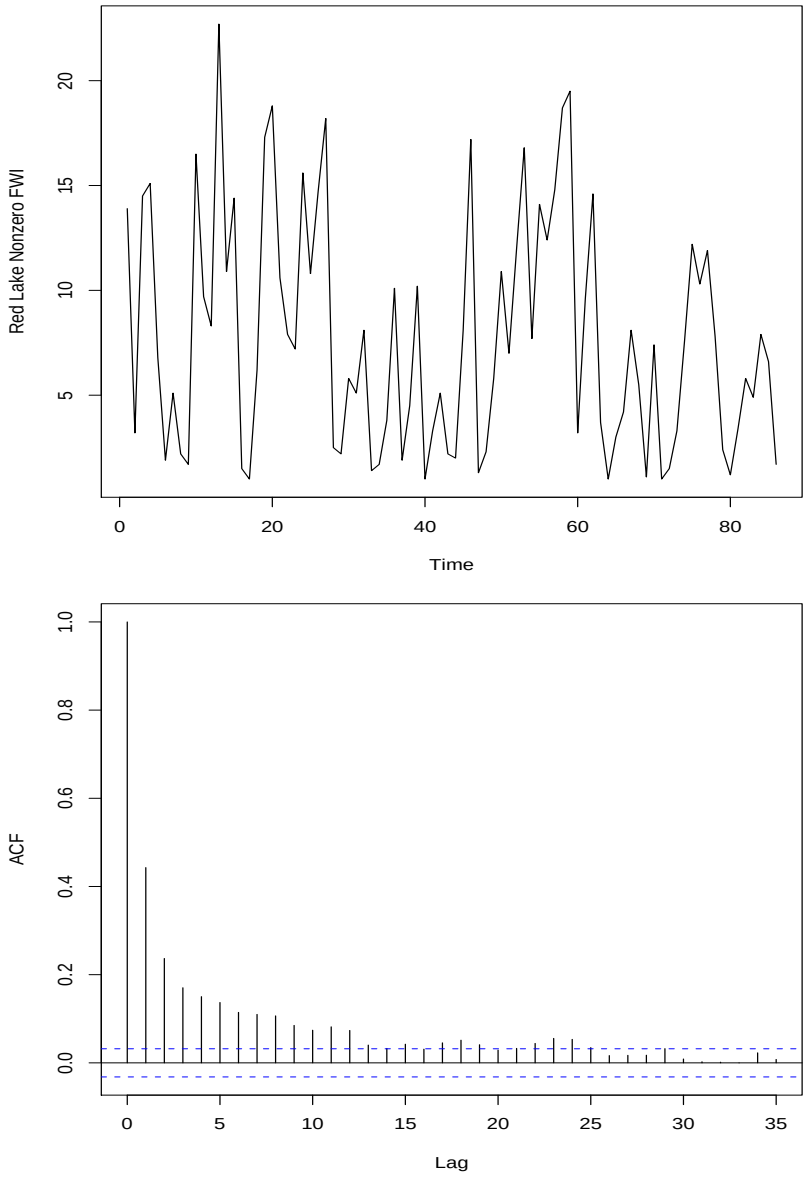

Figure 2.2: Upper Panel: Non-zero FWI Trace Plot at Red Lake for Fire Season 2000; Lower Panel: Autocorrelation function for all non-zero FWI at Red Lake (1963-2004).

Chapter 3

MINIFICATION MODELS FOR FWI

3.1 Introduction

The Fire Weather Index (FWI) is an important measure of fire danger in Canada, often interpreted as a proxy for the amount of effort that would be required to suppress extant wildfires. This index is a principal summary measure in the Canadian Forest Fire Behaviour Prediction System (Forestry Canada, 1992). Each day, during the Canadian fire season, the FWI is calculated using current and recent precipitation, temperature, relative humidity and wind speed measurements. These measurements are recorded at noon each day at several hundred weather stations located across the country. Our focus will be on the time-evolution behaviour of the FWI at six weather stations in the Province of Ontario, as test cases for a FWI time series model which we will propose in this paper.

The FWI was the subject of an important study by Martell (1999) in which 5-state Markov chains were employed as models for the evolution of the FWI at a number of weather stations in Ontario. Such models are governed by a large number of parameters, most of which are not easily interpretable. Thus, as pointed out in the paper itself, time series models for the FWI would be preferable. Such models usually have fewer parameters, and these parameters may be subject to clearer interpretation. Another advantage of time series models is that they more clearly reflect the numeric character of the process that they are emulating. An appropriate time series model can be used to directly simulate realistic FWI time series, while a 5-state Markov chain cannot be used for such a purpose. Furthermore, specific probabilities of FWI exceedance can be calculated from a time series model, while the corresponding exceedance probabilities coming from a Markov chain are necessarily much coarser. As we will see, the calculation of such exceedance probabilities for runs of one or more days of large FWI measurements could be of benefit to wildfire managers in terms of short-term suppression resource allocation.

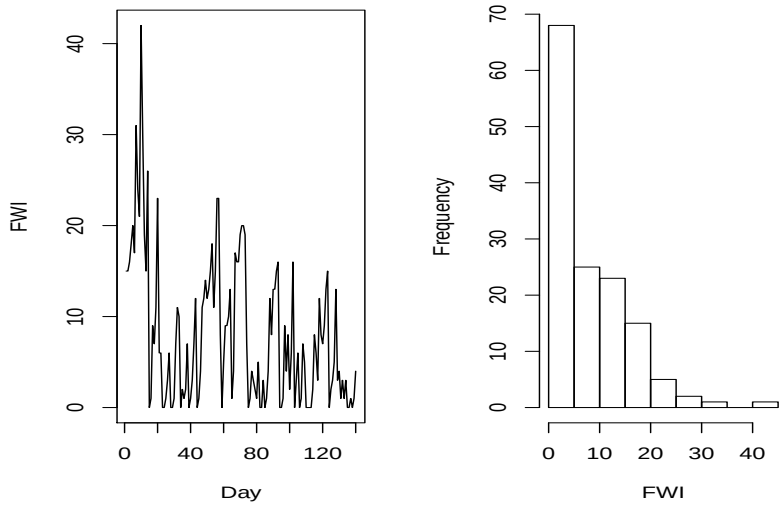

Figure 3.1: Trace plot and histogram of daily FWI measurements at Dryden, Ontario for April 29, 1987 through September 15, 1987.

3.1.1 Inadequacies of Linear Models for the FWI

Fujioka and Tsou (1985) developed first- and second-order autoregressive time series models for a related (American) fire weather index. Such models could provide quick rough approximations for the Canadian FWI, but accuracy may be questionable, as we now demonstrate.

A fairly typical situation is pictured in Figure 3.1: a trace plot of daily FWI values at Dryden, Ontario, for spring and summer, 1987 as well as the corresponding histogram. The histogram shows a high degree of skewness, so transformation of the data is required before applying an autoregressive model. Because of 0's in the data, we add 1 to all observations before applying a log transformation. For day I, set

$$Z_i = \log(\text{FWI}_i + 1).$$

The histogram in Figure 3.2 indicates that the log transformation has substantially reduced the skewness in the data, and the sample autocorrelation plot indicates that a first-order autoregressive process might provide an adequate model for the transformed data.

Using the `arima()` function in R (R Core Development Team, 2014), the fitted

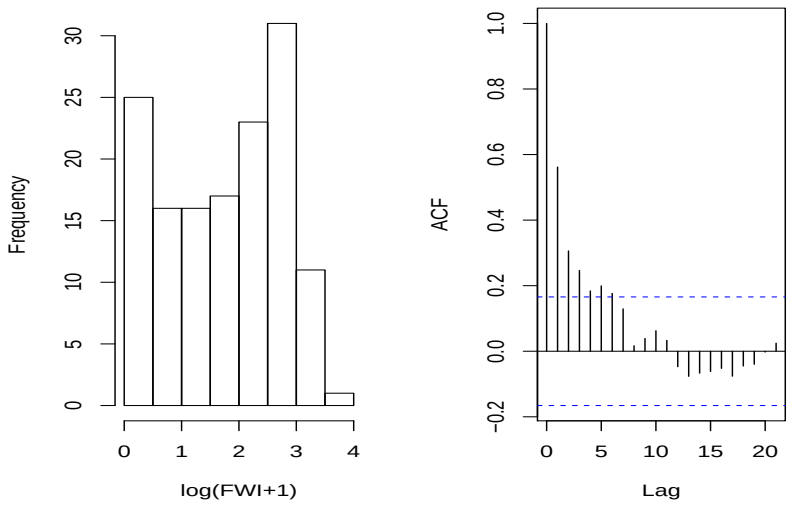

Figure 3.2: Histogram and sample autocorrelations of log-transformed Dryden FWI measurements.

model is found to be

$$Z_i = 1.7087 + 0.5620(Z_{i-1} - 1.7087) + \varepsilon_i$$

where ε_i is a random variable with mean 0 and variance 0.7998.

Standard model checks include a plot of the sample autocorrelations of the residuals. In this case, all such autocorrelations are very small, indicating a good model. A normal QQ-plot of the residuals indicates non-normality, as would be expected, but the model might be considered acceptable as an approximation, and for some purposes, this may be true. However, we will now show that the model cannot be used as a basis for simulating realistic FWI sequences.

Simulating from the above model is straightforward. Figure 3.3 shows trace plots of two simulation runs where the number of simulated observations matches the size of the 1987 Dryden sample.

The trace plot in the left panel of Figure 3.3 exhibits an FWI value near 70. This is extremely high; 40 years of observations at Dryden yield a maximum value of 60. The plot in the right panel is even more extreme: two simulated observations are in excess of 100, including one above 150. Such values are not expected under

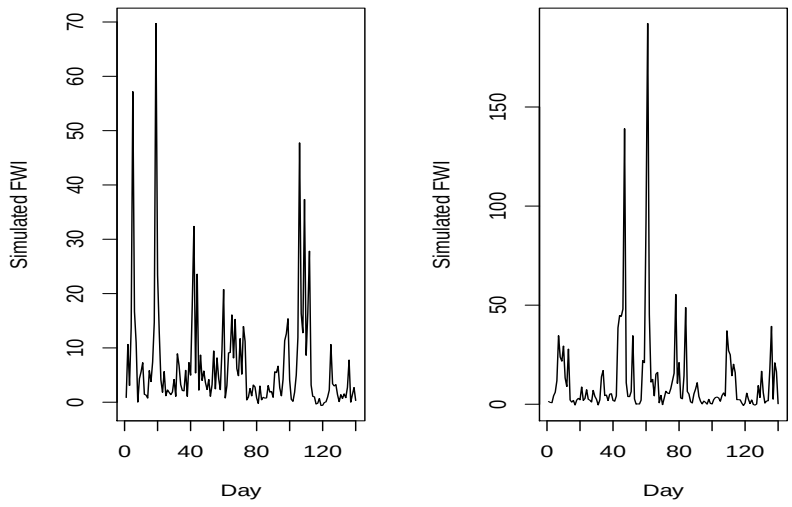

Figure 3.3: Trace plots for two runs of data, simulated from the fitted first order autoregressive model, back-transformed to the original scale.

any realistic scenario. Thus, a realistic FWI simulator must be based on an entirely different kind of time series model.

3.1.2 Model Considerations Based on Data Exploration

Returning to the trace plot in the left panel of Figure 3.1, we see that the FWI measurements are nonnegative, and we also see occasional runs of 0's. The histogram in the right panel indicates that the data are roughly exponentially distributed, though closer examination of these data as well as measurements from other years and other weather stations indicates that the tail behaviour may be somewhat lighter than predicted by an exponential model. The autocorrelation plot in Figure 3.2 shows short range dependence.

The preceding discussion suggests that the appropriate modelling strategy will be through the use of nongaussian time series, particularly nonnegative-valued series which include atoms at 0.

3.1.3 Some Nonnegative Time Series Models

3.1.3.1 Minification Models

A univariate time series $\{X_n\}$ follows a minification process when

$$X_n = k \min(X_{n-1}, E_n) \tag{3.1}$$

for an appropriate constant $k > 1$. The random innovation E_n is usually nonnegative and assumed to be independent of X_{n-1}.

When E_n has an exponential distribution, the model at (3.1) is called an exponential minification process; it was first studied by Tavares (1980) who exploited the minimum property of exponential distributions to show that the process gives rise to a stationary time series of exponential random variables, although the use of the minimization operation to produce multivariate exponential distributions originates with Marshall and Olkin (1967). Among other things, Tavares showed that if $\{E_n\}$ is an i.i.d. sequence of exponential random variables with rate $(k-1)\lambda$, then the process yields a stationary sequence of random variables $\{X_n\}$ whose marginal distributions are exponential with rate λ. He showed that the lag i autocorrelations for this process are k^{-i}, for $i = 1, 2, \ldots$.

Sim (1986) developed simulation algorithms for a version of the minification process where the random variables E_n follow a Weibull distribution. In other words, $Y_n = X_n^{\mathrm{P}}$ follows model (3.1) when E_n is Weibull with shape parameter P and a scale parameter which matches that of the exponential variable X_n. Adke and Balakrishna (1992) carried out statistical inference for the case where E_n is exponentially distributed.

Lewis and McKenzie (1991) studied the minification process for cases where the innovations are not necessarily exponential or Weibull. Balakrishna and Jacob (2003) showed that the general minification process is ergodic, and used this result to establish consistency of the estimators of the mean and k.

3.1.3.2 Related Models

Another important autocorrelated process with exponential marginal distributions is the first-order exponential autoregressive process (EAR(1)) which is due to Gaver and

Lewis (1980):

$$X_n = \begin{cases} \rho X_{n-1}, & \text{w.p. } \rho. \\ \rho X_{n-1} + E_n, & \text{w.p. } 1-\rho. \end{cases} \tag{3.2}$$

where the E_i are independent exponential random variables with rate λ.

The EAR(1) process was shown by Chernick et al (1988) to be the time-reversed process of Tavares' (1980) exponential minification process.

We note that Lawrance and Lewis (1977) introduced a first-order exponential moving average model, EMA(1). Lawrance and Lewis (1980) also generalized the EAR(1) and EMA(1) to higher order and connected the two processes in the EARMA(p,q) model. Lawrance and Lewis (1981) introduced a generalization of the EAR(1) model which they called the new exponential autoregressive process, the NEAR(1) model:

$$X_n = \beta B_n X_{n-1} + \varepsilon_n \tag{3.3}$$

where B_n is a Bernoulli random variable with parameter α, and ε_n is a mixture of two exponential random variables: having rate λ, with probability $(1-\beta)(1-(1-\alpha)\beta)^{-1}$, and rate $\lambda((1-\alpha)\beta)^{-1}$, otherwise. The i.i.d sequences $\{B_n\}$ and $\{\varepsilon_n\}$ are independent. When $\alpha = 1$, the process is EAR(1).

The NEAR(1) model was designed to address some problems that were evident with the EAR(1) (as well as the associated minification) model, namely, that the processes are always positively autocorrelated and that runs of geometrically decreasing values occur. This last problem can be observed from (3.2), where if $X_1 = x$, then the event $\{X_2 = \rho x, X_3 = \rho^2 x, \ldots, X_{j+1} = \rho^j x\}$ occurs with probability $\rho^j > 0$.

The NEAR(1) model has exponential marginal distributions, but the evolution from X_{n-1} to X_n no longer includes geometrically changing runs.

The NEAR(1) process continues to have nonnegative autocorrelations. To obtain an exponential time series model with negative first lag autocorrelation, Lawrance and Lewis (1981) proposed a cross-coupled version of the NEAR(1) model using antithetic variable sampling. Specifically, sequences $\{X_n\}$ and $\{X'_n\}$ are defined as follows:

$$X_n = \beta B_n X'_{n-1} + \varepsilon_n$$

and

$$X'_n = \beta B'_n X_{n-1} + \varepsilon'_n$$

where the pairs (B_n', B_n) and $(\varepsilon_n', \varepsilon_n)$ are both negatively related. The marginal distributions of X_n and X_n' are both exponential, but the lag 1 autocorrelations within each sequence are negative.

3.1.3.3 Tailed Time Series Models

The models discussed to this point give rise to purely positive time series, if the innovations are always positive. Note that if the innovations are 0 with nonzero probability, there is a nonzero probability that these processes will be absorbed at 0 in finite time. In order to properly model the Fire Weather Index, atoms at 0 are needed, but the process should not be absorbed into the 0 state.

One possibility is to use a tailed model which is a mixture of a point mass at 0 with a continuous distribution on the positive halfline. The terminology "tailed model" is due to Littlejohn (1994) who introduced an exponential tailed minification process:

$$X_n = \min(X_{n-1}, E_n)/\rho + \begin{cases} \gamma_n, & \text{w.p. } \rho, \text{ if } \min(X_{n-1}, E_n) = 0. \\ 0, & \text{otherwise.} \end{cases}$$

where

$$\begin{aligned} P(E_n \geq x) &= (1-\phi)e^{-\frac{x\lambda}{\rho}+x\lambda}, \\ \phi &= \frac{\theta(1-\rho)}{\rho(1-\theta)} \end{aligned}$$

and

$$P(\gamma_n > x) = e^{-\lambda x}.$$

When $\rho = 1-\theta$, $\{X_n\}$ is a stationary exponential tailed process.

3.1.4 Outline of Paper

In Section 2, we consider the physical processes underlying the FWI in order to set out a basic time series model, and this is checked against the 1987 Dryden FWI data. A step-by-step approach is taken in Sections 2 through 5 to gradually add features to the time series model with the goal of matching the characteristics of the data more closely. At each stage, the current model is checked by comparing a trace plot and an autocorrelation plot with the plots given in Figure (3.2). We note that matching the autocorrelation function for a nonlinear time series model with the sample autocorrelation function is not sufficient to ensure model adequacy, but a failure to match is an indicator of model inadequacy.

In Section 5, a random coefficient model will be described and studied, both at a general level and at a level specific enough to be fit to our data. In Section 6, we apply the random coefficient model to the entire set of data from six weather stations, while considering additional ways to check model adequacy. Applications of the random coefficient minification model are considered in Section 7. Concluding remarks can be found in the final section.

3.2 FWI and Exponential Tailed Minification Processes

The FWI is calculated from the fine fuel moisture code (FFMC), wind speed and the Buildup Index (BUI). The FFMC, in turn, is governed by temperature, relative humidity, wind speed and precipitation, while the BUI is dominated by precipitation, with some temperature and relative humidity effects through the dependence of the BUI on the duff moisture code (DMC). Thus, drying and wetting through atmospheric processes play a large role in the FWI, with precipitation leading to occasional abrupt changes (often, but not always, to nil values).

In the Canadian Forest Fire Danger Rating System (Van Wagner, 1987), wetting and drying towards equilibrium moisture values are described, particularly for the FFMC. Because of its high degree of dependence on the FFMC, the FWI might also be thought to have "equilibrium" values which it gravitates towards. Therefore, when "substantial" precipitation effects are absent, we expect the FWI to gradually increase or decrease depending upon which side of the "equilibrium" it is. However, precipitation effects are somewhat likely to cause arbitrarily-sized decreases, though never to negative values. Thus, it seems reasonable to suppose that drying effects from day to day, when wetter than the equilibrium value, will be essentially constant, while

wetting effects, even when drier than equilibrium, will be somewhat less predictable. A basic model which has this kind of character is a minification model:

$$X_n = (\alpha + 1)\min(X_{n-1}, \alpha^{-1}\varepsilon_n) \tag{3.4}$$

where ε_n is independent of X_{n-1} and has a distribution which is, in a vague sense, centered at or near the FWI equilibrium value. The sequence of X's coming from this model will tend to gradually increase towards the equilibrium from below and to somewhat unpredictably drop back below the equilibrium value after a period of increase.

3.2.1 A Basic Model

To more fully model precipitation effects, we need to include a mixture component at 0. One possibility is

$$X_n = \begin{cases} (\alpha + 1)\min(X_{n-1}, \alpha^{-1}\varepsilon_n), & \text{if } X_{n-1} > 0. \\ \varepsilon_n, & \text{if } X_{n-1} = 0. \end{cases} \tag{3.5}$$

where ε_n has a tailed model which is a mixture of a point mass at 0 with a continuous distribution on the positive halfline. For our present purpose, we assume that ε_n has an exponential tailed distribution with parameters λ and p_ε ($\mathrm{ET}(\lambda, p_\varepsilon)$) random variable if

$$P(\varepsilon_n \geq x) = (1 - p_\varepsilon)e^{-\lambda x},$$

where $p_\varepsilon \in (0, 1)$.

3.2.1.1 Stationarity

It is possible to define a stationary process via (3.5). To see this directly, first note that

$$\begin{aligned} P(X_n = 0) &= P(X_n = 0|X_{n-1} > 0)P(X_{n-1} > 0) + \\ &\quad P(X_n = 0|X_{n-1} = 0)P(X_{n-1} = 0) \\ &= p_\varepsilon(1 - P(X_{n-1} = 0)) + p_\varepsilon P(X_{n-1} = 0) \\ &= p_\varepsilon. \end{aligned}$$

Now, suppose that

$$P(X_{n-1} \geq x) = (1 - p_\varepsilon)e^{-\lambda x}.$$

It follows that

$$\begin{aligned} P(X_n \geq x) &= p_\varepsilon(1 - p_\varepsilon)e^{-\lambda x} + \\ &\quad P\left(X_{n-1} \geq x(\alpha+1)^{-1}\varepsilon_n \geq x\alpha(\alpha+1)^{-1}|X_{n-1} > 0\right) P(X_{n-1} > 0) \\ &= p_\varepsilon(1 - p_\varepsilon)e^{-\lambda x} + (1 - p_\varepsilon)(1 - p_\varepsilon)e^{-\lambda x} = (1 - p_\varepsilon)e^{-\lambda x}. \end{aligned}$$

Therefore, $\{X_n\}$ is a stationary exponential tailed process. The model for FWI at (3.5) is essentially equivalent to the Littlejohn (1994) model for Y_n.

3.2.1.2 Parameter Estimation

For the exponential tailed case of the minification model (3.5), there are 3 parameters: λ, p_ε and α. Results from Balakrishna and Jacob (2003) can be used to show that consistent parameter estimators using data of the form $X_1, X_2, \ldots, X_n$ are as follows:

$$\widehat{\alpha} = \max_{X_j > 0} \left(\frac{X_j}{X_{j-1}}\right) - 1$$

and

$$\widehat{p_\varepsilon} = \frac{1}{n}\sum_{j=1}^{n} I_{X_j = 0}$$

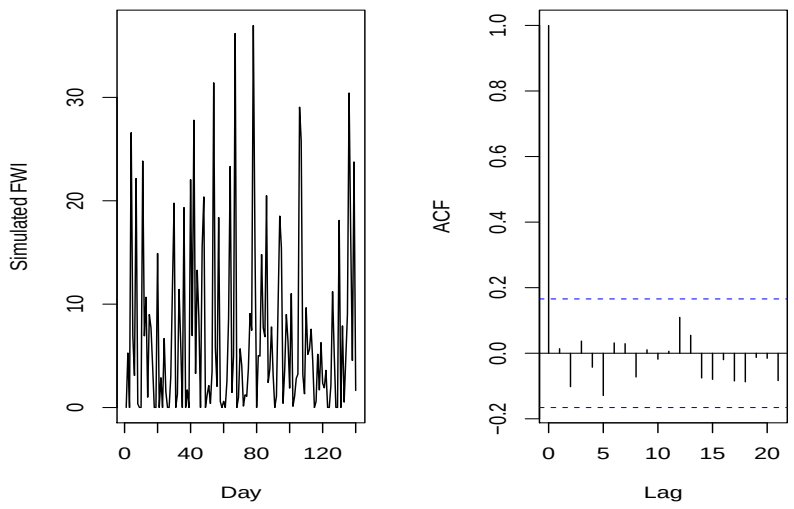

Figure 3.4: Left panel: simulated run of exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the year 1987. Right panel: autocorrelation estimates for the simulated data.

where $I_{X=0}$ denotes the indicator that $X = 0$. Since

$$P(\varepsilon_j = 0) = P(X_j = 0)$$

for this model, the rate parameter can be estimated using

$$\widehat{\lambda} = \frac{1 - \widehat{p_\varepsilon}}{\bar{X}}$$

where $\bar{X}$ denotes the average of the sample.

Parameter estimates for the exponential tailed minification model fit to the 1987 Dryden data are $\widehat{p_\varepsilon} = 0.179$, $\widehat{\alpha} = 8$ and $\widehat{\lambda} = 0.104$.

3.2.1.3 Model Checking

We can use simulation again to show that this model is not adequate for the data. In Figure 3.4, a trace plot and sample autocorrelation plot are displayed for data that have been simulated from model (3.5).

We observe that, although the simulated process is providing more realistically valued measurements, the autocorrelations for the simulated data are all small, compared with the relatively large lag 1 autocorrelation for the observed data.

3.2.1.4 Seasonality

Martell's (1999) work demonstrates that seasonality plays a role in the behaviour of the FWI in Ontario. He divided the fire season into 3 subseasons: April 29-June 9, June 10-July 28, and July 29-September 15. He then estimated his 5-state Markov chain transition parameters for each subseason.

Adopting a similar strategy, we can estimate the exponential tailed minification model parameters for each of Martell's subseasons. Parameter estimates for the seasonal version of the exponential tailed minification model fit to the 1987 Dryden data are given in Table 3.1.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.14	0.20	0.08
λ	0.09	0.10	0.10
α	9.00	7.00	2.50

Table 3.1: Parameter estimates for the exponential tailed model fit to each of the 3 subseasons of the 1987 Dryden FWI data.

In Table 3.1, p_ε appears to change from subseason to subseason. This could be due to different rainfall patterns in the different subseasons; this seems especially noticeable in the third subseason. The parameter α also appears to change. This parameter is connected to the rate of drying. The estimates of λ appear to be stable over the 3 subseasons, indicating that the baseline or mean FWI is essentially constant over the fire season. The conclusions we have drawn at this point are tentative and subject to further model validation as well as the study of additional data, to be carried out later.

The trace plot and autocorrelation plot for data simulated from the seasonal fitted exponential tailed minification model are displayed in Figure 3.5. Seasonality gives an improved fit, but there are still problems. The lag 1 autocorrelation is now clearly nonzero, but it is still much less than for the observed data. The autocorrelations at higher lags do not match those for the observed data at all. The trace plot indicates a tendency for larger values of the FWI to be generated than we see in the trace plot for the observed data.

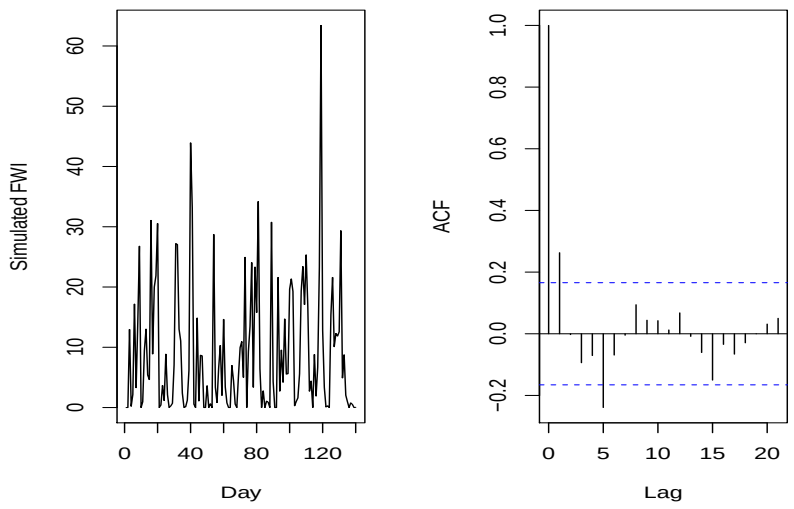

Figure 3.5: Left panel: simulated run of exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

There also may not be enough 0's in the simulated data; the process does not seem to stay at 0 long enough. This suggests changing the probability of transition from the 0 state from p_ε to something else.

3.2.2 A Modified Exponential Tailed Minification Model

More specifically, assume that $\{\delta_n\}$ is a sequence of independent $\mathrm{ET}(\lambda, p_\delta)$ random variables which are independent of $\{\varepsilon_n\}$. Define

$$X_n = \begin{cases} (\alpha+1)\min(X_{n-1}, \alpha^{-1}\varepsilon_n), & \text{if } X_{n-1} > 0. \\ \delta_n, & \text{if } X_{n-1} = 0. \end{cases} \tag{3.6}$$

3.2.2.1 Stationarity

It is still possible to define a stationary process via (3.6). To see this, set $p_{\mathrm{x}} = P(X_n = 0) = P(X_{n-1} = 0) = p_\varepsilon(1 + p_\varepsilon - p_\delta)^{-1}$, and suppose that

$$P(X_{n-1} \geq y) = (1 - p_{\mathrm{x}})e^{-\lambda y}.$$

It follows that

$$\begin{aligned} P(X_n \geq y) &= p_{\mathrm{x}}(1-p_\delta)e^{-\lambda y} + \\ &\quad P\left(X_{n-1} \geq y(\alpha+1)^{-1}, \varepsilon_n \geq y\lambda(\alpha+1)^{-1} | X_{n-1} > 0\right) P(X_{n-1} > 0) \\ &= p_{\mathrm{x}}(1-p_\delta)e^{-\lambda y} + (1-p_{\mathrm{x}})(1-p_\varepsilon)e^{-\lambda y} \\ &= (1-p_{\mathrm{x}})e^{-\lambda y}. \end{aligned}$$

Thus, $\{X_n\}$ is a stationary exponential tailed process.

3.2.2.2 Parameter Estimation

Consistent parameter estimators using data of the form $X_1, X_2, \ldots, X_n$ are as follows:

$$\widehat{\alpha} = \max_{X_j > 0} \left(\frac{X_j}{X_{j-1}} \right) - 1. \tag{3.7}$$

and

$$\widehat{p_\varepsilon} = \frac{1}{n_+} \sum_{j=2}^{n} I_{X_j=0, X_{j-1}>0} \tag{3.8}$$

where n_+ denotes the number of observations in $\{X_1, \ldots, X_{n-1}\}$ which are positive, since

$$P(\varepsilon_j = 0) = P(X_j = 0 | X_{j-1} > 0)$$

for this model.

$$\widehat{p_\delta} = \frac{1}{n_0} \sum_{j=2}^{n} I_{X_j=0, X_{j-1}=0} \tag{3.9}$$

where n_0 denotes the number of observations in $\{X_1, \ldots, X_{n-1}\}$ which are 0, since

$$P(\delta_j = 0) = P(X_j = 0 | X_{j-1} = 0)$$

for this model.

$$\widehat{\lambda} = \frac{1 - \widehat{p_{\mathrm{x}}}}{\bar{X}} \tag{3.10}$$

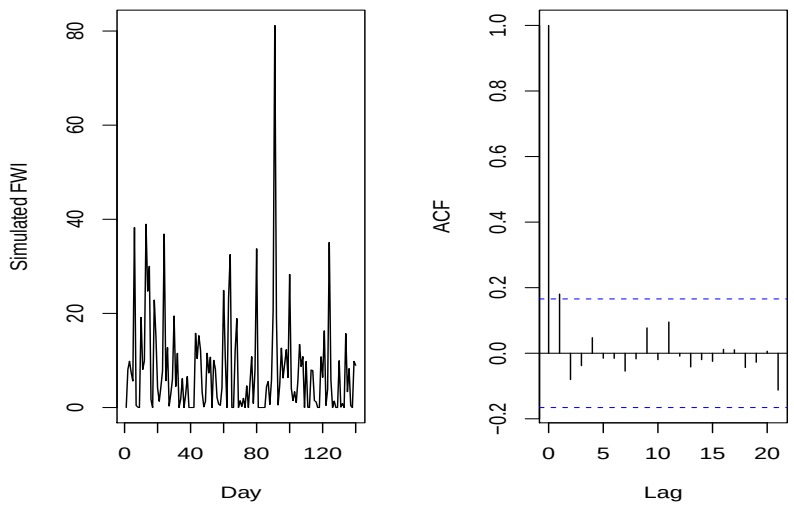

Figure 3.6: Left panel: simulated run of the modified exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the year 1987. Right panel: autocorrelation estimates for the simulated data.

where $\widehat{p_{\mathrm{x}}}$ denotes the proportion of 0's in the sample.

Parameter estimates for this generalization of the minification model fit to the 1987 Dryden data are: $\widehat{p_\varepsilon} = 0.149$, $\widehat{\alpha} = 8$, $\widehat{\lambda} = 0.107$, and $\widehat{p_\delta} = 0.32$. We note that the estimate of p_δ is much larger than the estimate of p_ε. Thus, the probability of remaining in the nil state is larger than the probability of entering the nil state.

3.2.2.3 Model Checking

We can use simulation again to show that this model is not adequate for the data. In Figure 3.6, a trace plot and sample autocorrelation plot are displayed for data that have been simulated from model (3.6) using parameter values estimated from the 1987 Dryden FWI data.

We observe that short runs of 0's are somewhat more evident now and the first lag autocorrelation is clearly nonzero, but still much less than the observed first lag autocorrelation.

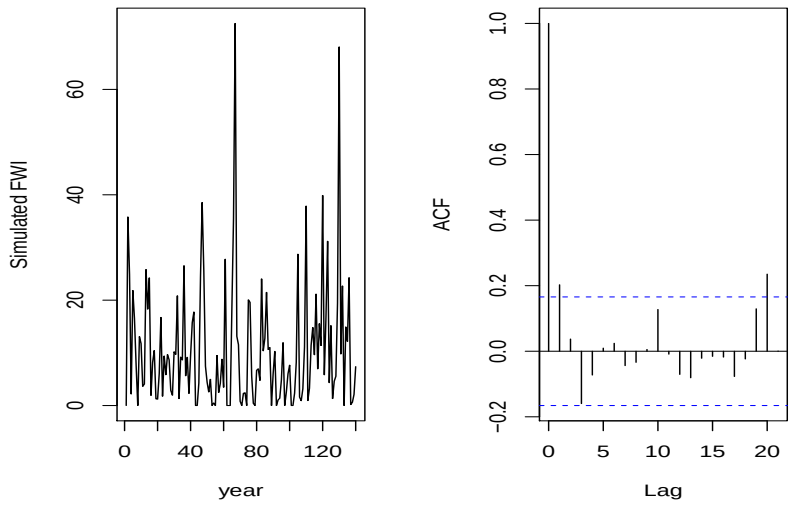

Figure 3.7: Left panel: simulated run of the modified exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

3.2.2.4 Seasonality

We can estimate the modified exponential tailed minification model parameters for each of Martell's (1999) subseasons. Parameter estimates for the seasonal version of the modified exponential tailed minification model fit to the 1987 Dryden data are given in Table 3.2.

The estimates of λ and α are based on the same estimators as the exponential tailed model so they remain unchanged. p_ε still appears to change from subseason to subseason, and p_δ follows the same general pattern but with correspondingly larger magnitudes. Any conclusions to be drawn remain tentative and subject to further model validation as well as the study of additional data.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.09	0.10	0.10
α	9.00	7.00	2.50
p_δ	0.33	0.40	0.25

Table 3.2: Parameter estimates for the modified exponential tailed model fit to each of the 3 subseasons of the 1987 Dryden FWI data.

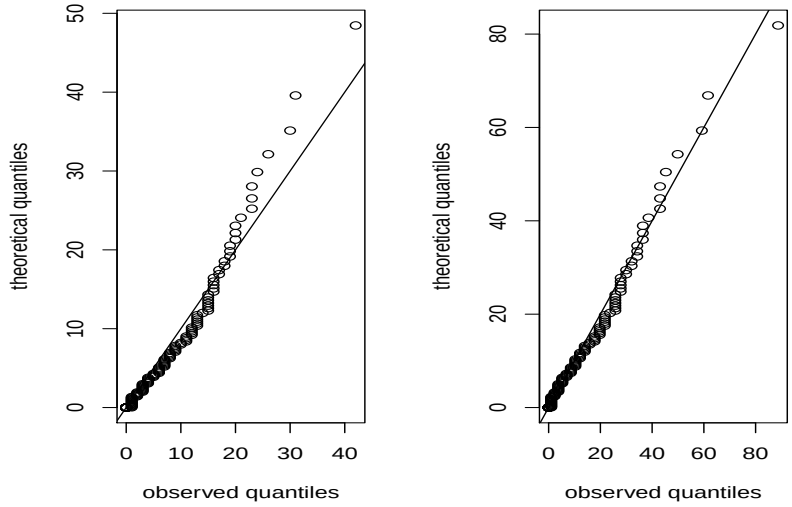

Figure 3.8: Left panel: Exponential tailed QQ-plot for FWI observations at Dryden, Ontario, for the year 1987. Right panel: Exponential tailed QQ-plot for the same observations raised to the power 1.2.

3.3 A Weibull Tailed Minification Model

3.3.1 Checking the Exponential Tailed Assumption

We can use a QQ-plot to see if the exponential tailed assumption is reasonable and to determine whether a power transformation will correct the problem. A power transformation corresponds to a Weibull tailed model. The qth quantile of the ET(λ, p_{X}) model is

$$x = \frac{\log(1 - p_{\mathrm{X}}) - \log(1 - q)}{\lambda}$$

unless $\log(1 - p_{\mathrm{X}}) - \log(1 - q) < 0$, in which case $x = 0$.

Figure 3.8 displays QQ-plots of the 1987 Dryden FWI data. The plot in the left panel is for the untransformed data and the plot in the right panel is of the data raised to the power 1.2. We observe that the untransformed data deviate from the 45° reference line. Therefore we conclude that transformation is needed. The transformed data lie close to the reference line indicating that the required power is near 1.2.

3.3.2 Details of the Model

We continue to assume that $\{\delta_n\}$ is a sequence of independent $\mathrm{ET}(\lambda, p_\delta)$ random variables which are independent of the $\mathrm{ET}(\lambda, p_\varepsilon)$ variables $\{\varepsilon_n\}$. Define

$$Y_n = X_n^{\mathrm{P}}$$

for some $\mathrm{P} > 0$ where X_n follows the model (3.6). That is,

$$Y_n = \begin{cases} (\alpha+1)^{\mathrm{P}} \min(Y_{n-1}, \alpha^{-\mathrm{P}} \varepsilon_n^{\mathrm{P}}), & \text{if } Y_{n-1} > 0. \\ \delta_n^{\mathrm{P}}, & \text{if } Y_{n-1} = 0. \end{cases} \tag{3.11}$$

Y_n is said to be Weibull tailed: $\mathrm{WT}(\lambda, p_{\mathrm{Y}}, \mathrm{P})$. It is easy to see that

$$p_{\mathrm{Y}} = p_{\mathrm{X}} = \frac{p_\varepsilon}{1 - p_\delta + p_\varepsilon}.$$

The right panel of Figure 3.8 suggests that $\mathrm{FWI}^{1.2}$ is $\mathrm{ET}(\lambda, p_{\mathrm{X}})$. Thus, $Y = \mathrm{FWI} \sim \mathrm{WT}(\lambda, p_{\mathrm{X}}, 5/6)$.

3.3.2.1 Stationarity

It is still possible to define a stationary process via (3.11). To see this, set

$$P(Y_{n-1} \geq y) = (1 - p_{\mathrm{X}}) e^{-\lambda y^{1/\mathrm{P}}}.$$

It follows that

$$P(Y_n \geq y) = P(X_n \geq y^{1/\mathrm{P}}) = (1 - p_{\mathrm{X}}) e^{-\lambda y^{1/\mathrm{P}}}.$$

Therefore, $\{Y_n\}$ is a stationary Weibull tailed process.

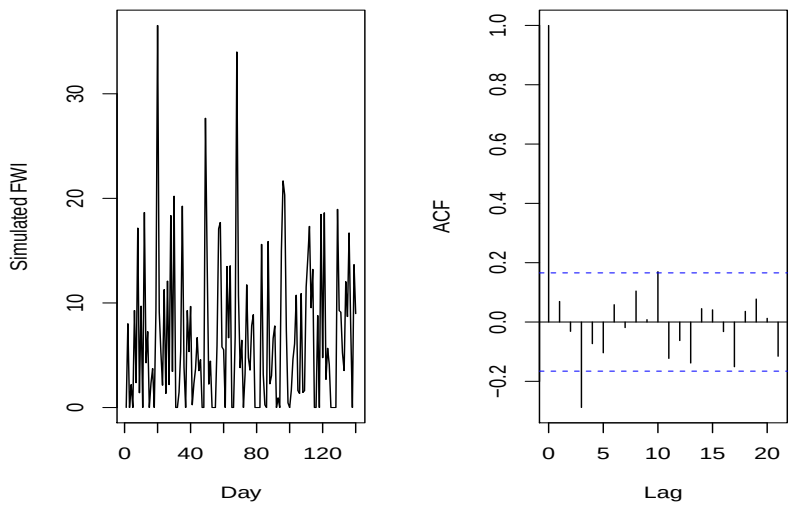

Figure 3.9: Left panel: simulated run of the modified Weibull tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the year 1987. Right panel: autocorrelation estimates for the simulated data.

3.3.2.2 Parameter Estimation

Given P, consistent parameter estimators using data of the form $Y_1, Y_2, \ldots, Y_n$ are given by equations (3.7) through (3.10) with $X_j = Y_j^{1/\mathrm{P}}$, for $j = 1, \ldots, n$. Parameter estimates for the Weibull minification model fit to the 1987 Dryden data, using P = .833, are: $\widehat{p_\varepsilon} = 0.149$, $\widehat{\alpha} = 12.97$, $\widehat{\lambda} = 0.0637$, and $\widehat{p_\delta} = 0.32$.

3.3.2.3 Model Checking

We again use simulation to check adequacy of the fitted Weibull model. In Figure 3.9, a trace plot and sample autocorrelation plot are displayed for data that have been simulated from model (3.11). The autocorrelation plot indicates that the Weibull model is not capturing the dependence in the data properly. This is not surprising in view of our earlier observation that seasonality is playing a role.

3.3.2.4 Seasonality

Parameter estimates for the seasonal version of the modified Weibull tailed minification model fit to the 1987 Dryden data are given in Table 3.3.

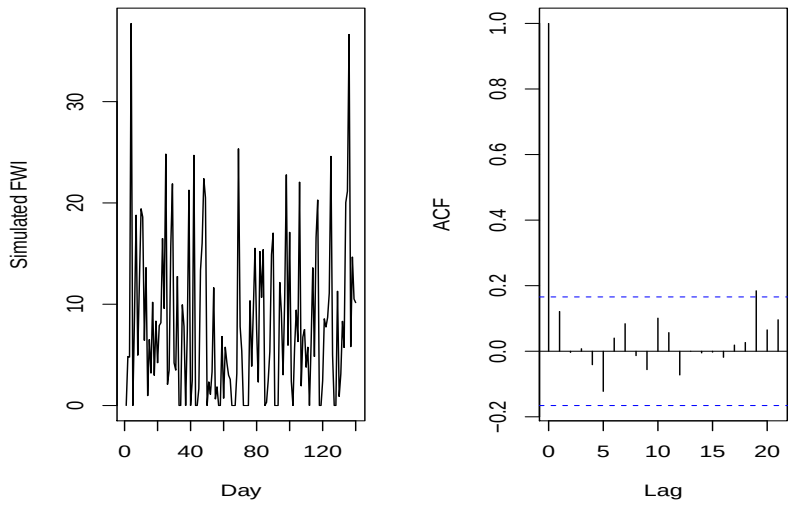

Figure 3.10: Left panel: simulated run of the modified Weibulll tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

The estimates of λ appear to be stable across the subseasons, and the estimates of α and p_ε still appear to change from subseason to subseason. Estimates of p_δ follow the same general pattern but with correspondingly larger magnitudes.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.05	0.06	0.06
α	14.85	11.13	3.50
p_δ	0.33	0.40	0.25

Table 3.3: Parameter estimates for the modified Weibull tailed model fit to each of the 3 subseasons of the 1987 Dryden FWI data, assuming P = 5/6.

In Figure 3.10, a trace plot and sample autocorrelation plot are displayed for data simulated from model (3.11) using parameters estimated for each of the subseasons using parameters estimated for each of the subseasons using the 1987 Dryden FWI data. The trace plot looks somewhat realistic, but the lag 1 autocorrelation is still disappointingly small. The model is still not satisfactory.

3.4 A Mixture Tailed Minification Process

3.4.1 Comparing the Distributions of δ and X

Under the assumptions of a Weibull tailed minification model, the marginal distributions of the continuous parts of δ_j and X_j are both Weibull tailed with parameters P and λ. Our present goal is to check whether the rate parameters for δ_j and ε_j are the same using the 1987 Dryden FWI data.

We can observe δ_j directly, since $X_j = \delta_j$, when $X_{j-1} = 0$. If the rate parameter for ε_n is the same, then the marginal distribution of X_j, given $X_{j-1} > 0$ and $X_j > 0$ should be the same as for ε_j. Therefore, we can compare the marginal distributions of δ_j and ε_j using a QQ-plot of these quantities.

Figure 3.11 displays such a QQ-plot for the FWI observations for Dryden in 1987 with a 45° reference line. There is a clear discrepancy between the observations and the reference line, indicating that the two distributions are very different. As a further confirmation, we have simulated five datasets from the fitted model and recomputed the QQ-plot in each case. In order for the reference plots to be completely comparable with the QQ-plot of the original data, we have applied the ceiling to all observations before sorting.

Although some of the reference plots display some departure from the reference line as well, none is as extreme as the plot of the original data. Also, the scale of the original dataset is much smaller than for any of the reference datasets. Our conclusion is that we should allow δ_j to have a different rate parameter from ε_j.

3.4.2 The Mixture Tailed Process

We now consider a version of the tailed model (3.6) where ε_n is ET(λ, p_ε) and δ_n is ET(γ, p_δ), where it is possible for γ and λ to be unequal. This leads to a stationary distribution which is a mixture of exponentials (or Weibulls).

3.4.2.1 Stationary Distribution

A stationary process can still be defined, but the marginal distribution is no longer exponential tailed. Instead, the continuous part of the distribution is a mixture

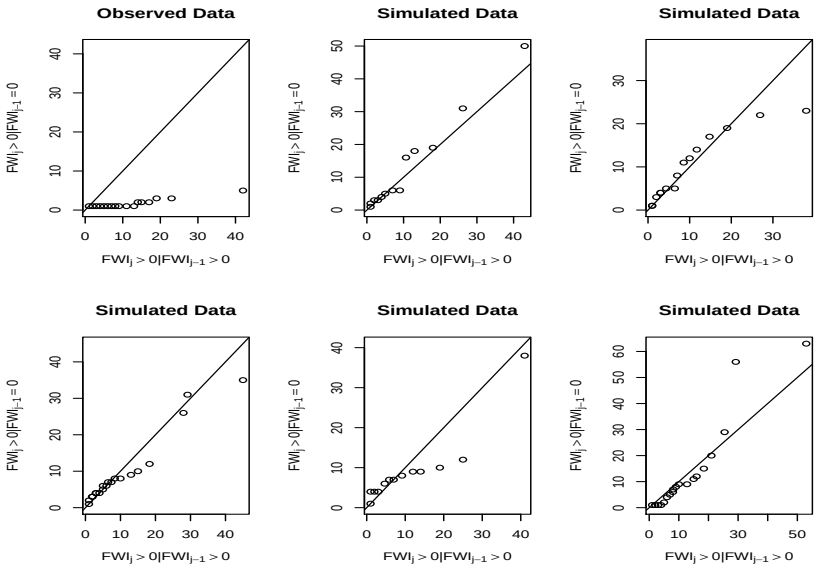

Figure 3.11: QQ-plots of δ_j versus X_j for the Weibull tailed minification model applied to FWI observations at Dryden, Ontario, for the year 1987 and data simulated from the fitted model.

of exponential tailed distributions. We continue to make the same independence assumptions as before.

$$
\begin{aligned}
P(X_n \geq x) &= p_{\mathrm{x}} P(\delta_n \geq x \mid X_{n-1} = 0) + \\
&\quad (1 - p_{\mathrm{x}}) P((\alpha+1)\min(X_{n-1}, \alpha^{-1}\varepsilon_n) \geq x \mid X_{n-1} > 0) \\
&= (1 - p_{\mathrm{x}}) P(X_{n-1} \geq x(\alpha+1)^{-1}, \varepsilon_n \geq \alpha x(\alpha+1)^{-1} \mid X_{n-1} > 0) + \\
&\quad p_{\mathrm{x}}(1 - p_\delta) e^{-\gamma x} \\
&= (1 - p_{\mathrm{x}})(1 - p_\varepsilon) e^{-\lambda\alpha x(\alpha+1)^{-1}} P(X_{n-1} \geq x(\alpha+1)^{-1} \mid X_{n-1} > 0) + \\
&\quad p_{\mathrm{x}}(1 - p_\delta) e^{-\gamma x} \\
&= (1 - p_\varepsilon) e^{-\lambda\alpha x(\alpha+1)^{-1}} P(X_{n-1} \geq x(\alpha+1)^{-1}, X_{n-1} > 0) + \\
&\quad p_{\mathrm{x}}(1 - p_\delta) e^{-\gamma x} \\
&= (1 - p_\varepsilon) e^{-\lambda\alpha x(\alpha+1)^{-1}} P(X_{n-1} \geq x(\alpha+1)^{-1}) + p_{\mathrm{x}}(1 - p_\delta) e^{-\gamma x}.
\end{aligned}
$$

Stationarity of $\{X_n\}$ implies that

$$P(X_n \geq x) = (1-p_\delta)p_{\mathrm{X}}e^{-\gamma x} + (1-p_\varepsilon)e^{-\alpha\lambda x(\alpha+1)^{-1}}P(X_n \geq x(\alpha+1)^{-1}) \quad (3.12)$$

from which it is possible to derive the formula

$$P(X_n \geq x) = p_{\mathrm{X}}(1-p_\delta)e^{-\lambda x}\left[\sum_{k=1}^{\infty}(1-p_\varepsilon)^k e^{-\frac{(\gamma-\lambda)x}{(\alpha+1)^k}}\right] + p_{\mathrm{X}}(1-p_\delta)e^{-\gamma x}. \quad (3.13)$$

For the mixture Weibull tailed case, we can also obtain an explicit expression. Specifically, suppose $Y_n = X_n^{\mathrm{P}}$ for some $\mathrm{P} \neq 0$, and suppose ε_n and δ_n are nonnegative exponential tailed random variables with the same parameter P: $\varepsilon \sim \mathrm{ET}(\lambda, p_\varepsilon)$ and $\delta \sim \mathrm{ET}(\gamma, p_\delta)$. Then

$$\begin{aligned} P(Y_n \geq x) &= P(X_n^{\mathrm{P}} \geq x) = P(X_n \geq x^{1/\mathrm{P}}) \\ &= p_{\mathrm{X}}(1-p_\delta)e^{-\gamma x^{1/\mathrm{P}}} + p_{\mathrm{X}}(1-p_\delta)e^{-\lambda x^{1/\mathrm{P}}}\sum_{k=1}^{\infty}(1-p_\varepsilon)^k e^{-\frac{(\gamma-\lambda)x^{1/\mathrm{P}}}{(\alpha+1)^k}} \end{aligned}$$

3.4.2.2 Numerical Calculation of Stationary Probabilities

Approximate values of $P(X_n \geq x)$ are obtainable from (3.13) upon truncating the infinite sum, but (3.12) can also be used directly, since $P(X_n > 0)$ is known to be $1 - p_{\mathrm{X}}$. Thus, we may approximate $P(X_n > \epsilon)$ by $1 - p_{\mathrm{X}}$ for some small $\epsilon > 0$. Then $P(X_n \geq x)$ can be calculated for any positive x by recursion. This method can provide better accuracy than (3.13).

To demonstrate the recursion technique clearly, we supply the R code (R Core Development Team, 2014) which can be used to calculate the stationary probability $P(X \geq x)$ for a given set of parameter values:

```
> survProb<- function(x, p.delta, p.eps, lambda, gamma, alpha) {
+     pX <- p.eps/(1+p.eps-p.delta)
+     if (x < 0.1) {
```

```
+         SP <- 1-pX
+     } else {
+  SP <- survProb(x/(alpha+1), p.delta, p.eps, lambda, gamma,
+ alpha)
+     }
+     (1-p.delta)*exp(-gamma*x)*pX +
+                     (1-p.eps)*exp(-alpha*lambda*x/(alpha+1))*SP
+ }
```

An example of its use is

```
> survProb(12, .62, .18, .12, .31, 8)

[1] 0.1289089
```

This is the approximate value of $P(X \geq 12)$ when $p_\delta = .62$, $p_\varepsilon = .18$, $\lambda = .12$, $\gamma = .31$ and $\alpha = 8$.

When Y follows the mixture Weibull tailed process where $Y = X^{\mathrm{P}}$, the same function can be used, upon recalling that

$$P(Y \geq y) = P(X \geq y^{1/\mathrm{P}}).$$

When P = .9, and the other parameters are as above, we can calculate $P(Y \geq 10)$ as

```
> survProb(10^(1/.9), .62, .18, .12, .31, 8)

[1] 0.1147192
```

3.4.2.3 Fitting and Checking the Model

We consider the mixture exponential tailed case first. The estimators for α, p_ε and p_δ remain as given by equations (3.7), (3.8) and (3.9). The maximum likelihood estimator for γ is

$$\widehat{\gamma} = \frac{n_0}{\sum_{j=2}^{n} X_j I_{X_j>0, X_{j-1}=0}}. \tag{3.14}$$

In order to estimate λ, we first note that, provided that the sample is large enough, the maximum likelihood estimator for α is exact. This type of behaviour has been noted earlier (e.g. Gaver and Lewis, 1980). The likelihood is given by

$$L(\lambda, \alpha, p_\varepsilon, p_\delta) = \prod_{j=2}^{n} f_{\mathrm{x}_j|\mathrm{x}_{j-1}}(X_j \mid X_{j-1}) f_{\mathrm{x}_1}(X_1)$$

because the minification process is Markovian.

The conditional density of X_j, given the $X_{j-1} = 0$ is given by

$$f_{\mathrm{x}_j|\mathrm{x}_{j-1}}(X_j \mid 0) = (1-p_\delta)\gamma e^{-\gamma \mathrm{x}_j} + p_\delta I_{\{\mathrm{x}_j=0\}}$$

and is independent of λ. When $X_{j-1} > 0$, we have

$$\begin{aligned} f_{\mathrm{x}_j|\mathrm{x}_{j-1}}(X_j \mid X_{j-1}) &= P(\varepsilon_j \geq \alpha y) I_{\{\mathrm{x}_{j-1}=(\alpha+1)^{-1}\mathrm{x}_j\}} + \\ &\quad \alpha(\alpha+1)^{-1} f_\varepsilon(\alpha(\alpha+1)^{-1}X_j) I_{\{\mathrm{x}_{j-1}>(\alpha+1)^{-1}\mathrm{x}_j\}} \end{aligned}$$

since

$$\begin{aligned} P(X_j \geq x \mid X_{j-1} = y) &= P(X_{j-1} \geq x(\alpha+1)^{-1}, \varepsilon_j \geq \alpha x(\alpha+1)^{-1} \mid X_{j-1} = y) \\ &= P(\varepsilon_j \geq \alpha(\alpha+1)^{-1}x) I_{\{y \geq x(\alpha+1)^{-1}\}} \\ &= P(\varepsilon_j \geq \alpha y) I_{\{y = x(\alpha+1)^{-1}\}} + \\ &\quad P(\varepsilon_j \geq \alpha(\alpha+1)^{-1}x) I_{\{y > x(\alpha+1)^{-1}\}}. \end{aligned}$$

The expression to be maximized with respect to λ is thus

$$\begin{aligned} L(\lambda) &= \prod_{j:X_j=(\alpha+1)X_{j-1}} P(\varepsilon_j \geq \alpha X_{j-1}) \\ &\quad \prod_{k:X_k<X_{k-1}(\alpha+1)} f_\varepsilon(X_k\alpha(\alpha+1)^{-1})\alpha(\alpha+1)^{-1} \cdot f_{\mathrm{x}_1}(X_1) \end{aligned} \tag{3.15}$$

where the products are taken over j and k when $X_{j-1} > 0$ and $X_{K-1} > 0$. Equation (3.15) is equal to

$$\begin{aligned} L(\lambda) &= \prod_{j:X_j=(\alpha+1)X_{j-1}} (1-p_\varepsilon)e^{-\lambda\alpha X_{j-1}} \\ &\quad \prod_{k:X_k<X_{k-1}(\alpha+1)} \lambda(1-p_\varepsilon)e^{-\lambda X_k\alpha(\alpha+1)^{-1}}\alpha(\alpha+1)^{-1} \cdot f_{\mathrm{x}_1}(X_1) \end{aligned}$$

Maximizing this numerically gives the MLE for λ. A quick approximation is obtained by neglecting $f_{\mathrm{x}_1}(X_1)$. This will be justified only when the sample size large enough to trade X_1 as a missing observation. Proceeding this way, and upon differentiating the log likelihood with respect to λ we have

$$\alpha \sum_{j:X_j=(\alpha+1)X_{j-1}} X_{j-1} \sum_{k:X_k<X_{k-1}(\alpha+1)} (\lambda^{-1} - X_k\alpha(\alpha+1)^{-1}) = 0.$$

Letting n_ε denote the number of X_k's for which $X_k < (\alpha+1)X_{k-1}$, the approximate maximum likelihood estimator for λ is given by

$$\widehat{\lambda} = \frac{-n_\varepsilon}{\alpha\sum_{j:X_j=(\alpha+1)X_{j-1}} X_{j-1} - \sum_{k:X_k<X_{k-1}(\alpha+1)} X_k\alpha(\alpha+1)^{-1}}. \tag{3.16}$$

For the 1987 Dryden data, based on the approximate MLE method we found that $\widehat{\lambda} = 0.120$, $\widehat{p_\varepsilon} = 0.149$, $\widehat{p_\delta} = 0.32$, $\widehat{\gamma} = 0.607$ and $\widehat{\alpha} = 8$.

Figure 3.12 displays the trace plot for data simulated from the mixture exponential tailed minification model using parameters estimated by maximum likelihood

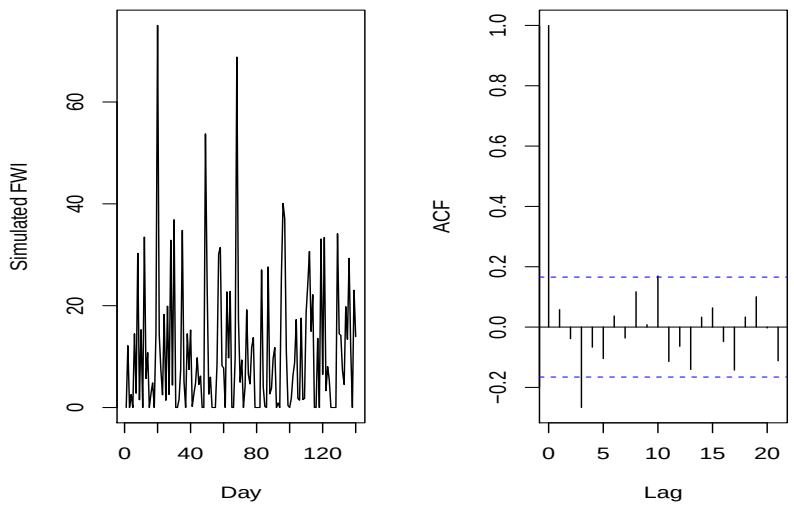

Figure 3.12: Left panel: simulated run of mixture exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season using maximum likelihood. Right panel: autocorrelation estimates for the simulated data.

on the 1987 Dryden FWI data. It can be seen that there is no real improvement; the trace plot shows several unrealistically large FWI values, while the autocorrelations are all small, with the exception of an anomalous negative value at lag 3.

The maximum likelihood estimator of λ depends very strongly on the correctness of the model. If the model is misspecified and α is incorrect, then λ could be badly biased. The next method provides an estimator which is not so heavily dependent on α. An alternative approach to parameter estimation in nonlinear time series is a least-squares method which minimizes the distance between the model-based conditional expectation $E[X_j \mid X_{j-1}]$ and the observation X_j:

For our purpose, we are interested in estimating λ so we solve the problem

$$\min_{\lambda} \sum_{j=2}^{n} (X_j - E[X_j \mid X_{j-1}])^2 \tag{3.17}$$

The first conditional lagged moment is

$$E[X_n | X_{n-1} = x] = (1 - p_\varepsilon)\lambda^{-1}(\alpha + 1)\alpha^{-1}(1 - e^{-\lambda x \alpha}).$$

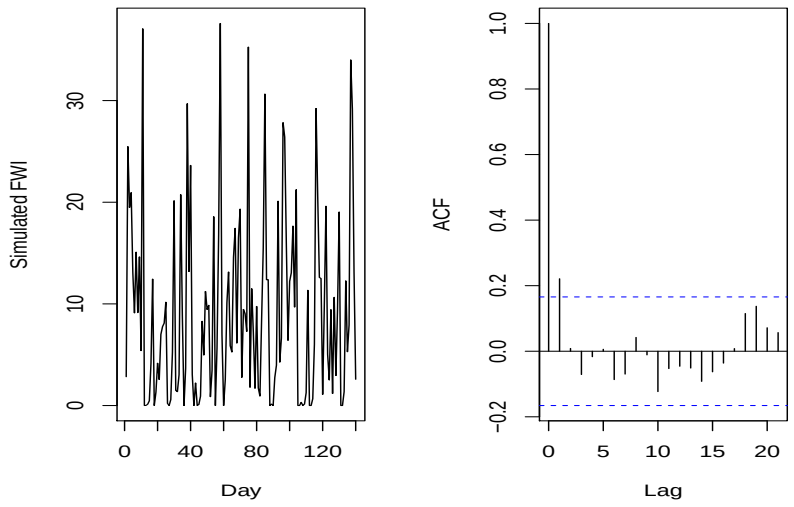

Figure 3.13: Left panel: simulated run of mixture exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season. The λ parameter has been estimated by least-squares. Right panel: autocorrelation estimates for the simulated data.

The `optim()` function in R (R Core Development Team 2014) can be used to find the solution to the equation (3.17).

For the 1987 Dryden data, we found that $\widehat{\lambda}$ changes to 0.087 when using the least-squares method. The other parameter estimates are the same as before.

Figure 3.13 contains a trace plot and autocorrelation plot for data simulated from the fitted model. The trace plot displays more realistic values than before, and the lag 1 autocorrelation is slightly larger. All other autocorrelations remain small. Thus, the model is inadequate whether the λ parameter is estimated by maximum likelihood or least-squares.

We next consider the Weibull case. When the FWI is transformed as $\text{FWI}^{6/5}$, the parameter estimates for p_ε and p_δ are the same as for the exponential case, while the maximum likelihood estimates are $\widehat{\lambda} = 0.073$, $\widehat{\alpha} = 12.967$ and $\widehat{\gamma} = 0.607$. When least-squares is used, then $\widehat{\lambda} = 0.049$.

A trace plot and autocorrelation plot for data simulated from the fitted mixture Weibull tailed minification model are given in Figure 3.14. They are similar in pattern to the exponential case. The model remains insufficient for the data.

When considering seasonality for the mixture exponential case, the estimates are shown in Table 3.4.

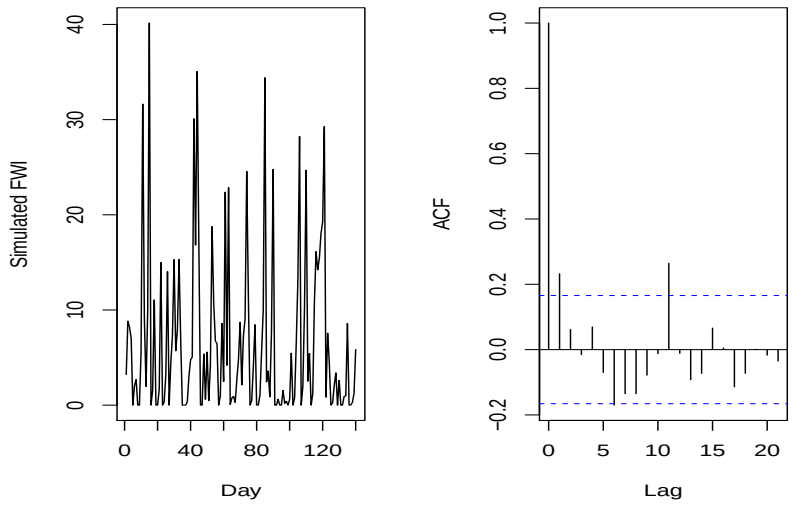

Figure 3.14: Left panel: simulated run of mixture Weibull tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, the 1987 fire season. The estimate of λ was obtained by least squares, and the parameter P is assumed to be 5/6. Right panel: autocorrelation estimates for the simulated data.

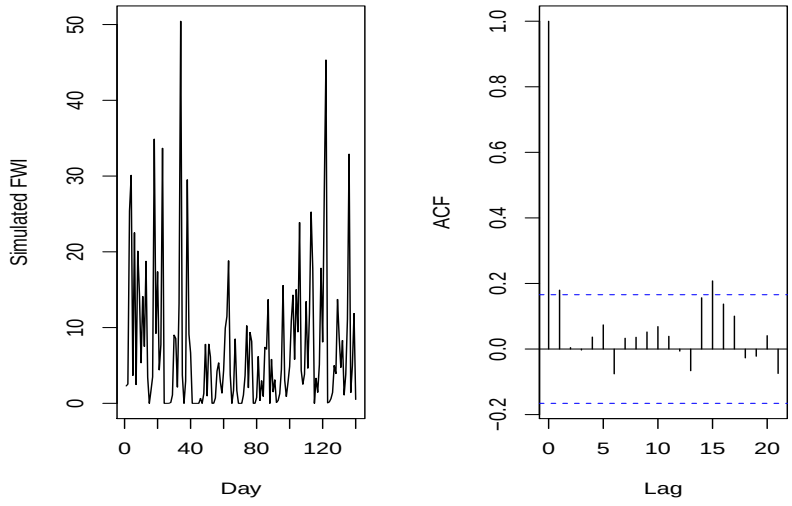

Figure 3.15: Left panel: simulated run of mixture exponential tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. The estimate of λ was obtained by least squares. Right panel: autocorrelation estimates for the simulated data.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.08	0.09	0.11
α	9.00	7.00	2.50
p_δ	0.33	0.40	0.25
γ	0.67	0.67	0.67

Table 3.4: Parameter estimates for the exponential tailed model fit to each of the 3 subseasons of the 1987 Dryden FWI data.The estimates of λ were obtained by least-squares.

Again over the three subseasons, λ seems stable, p_ε appears to change and estimates of α are the same as for the exponential tailed model. Estimates of γ for the mixture exponential tailed model are the same in all subseasons. p_δ appears to change and is larger than p_ε.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.04	0.05	0.06
α	14.85	11.13	3.50
p_δ	0.33	0.40	0.25
γ	0.59	0.59	0.59

Table 3.5: Parameter estimates for the mixture Weibull tailed minification model fit to each of the 3 subseasons of the 1987 Dryden FWI data. The estimates of λ were obtained by least-squares

Estimates of seasonal effects for the mixture Weibull tailed case are in Table 3.5 where P is taken to be 5/6. Estimates of λ seem stable over the three subseasons. p_ε and p_δ are the same as before transformation. Estimates of γ remain the same across the subseasons. As in the case of the untransformed data, α seems to decrease across the subseasons.

Figure 3.17 contains the trace and autocorrelation plots for the fitted seasonal mixture Weibull tailed process. The trace plot displays two very large values (over 40), which is not realistic. The low lag autocorrelations are higher than for most other models we have considered, but still much lower than for the observed data. The mixture minification model is an improvement, but it is still not an adequate model for the observed FWI data.

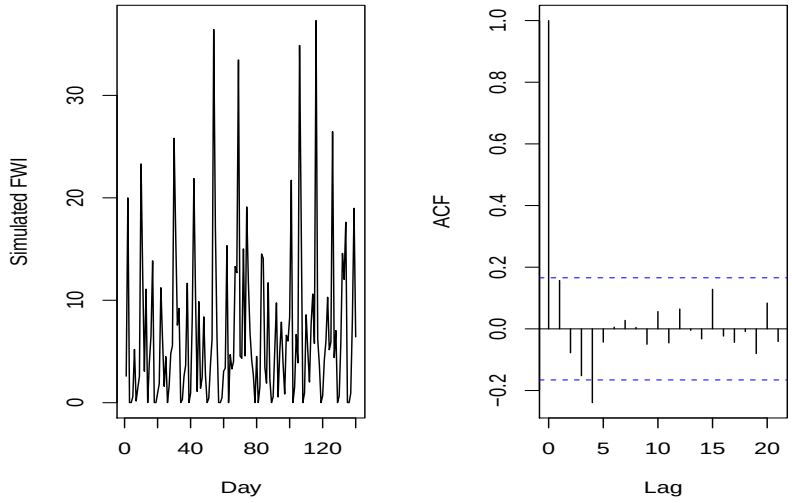

Figure 3.16: Left panel: simulated run of mixture Weibull tailed minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season. The parameter λ was estimated by least-squares. Right panel: autocorrelation estimates for the simulated data.

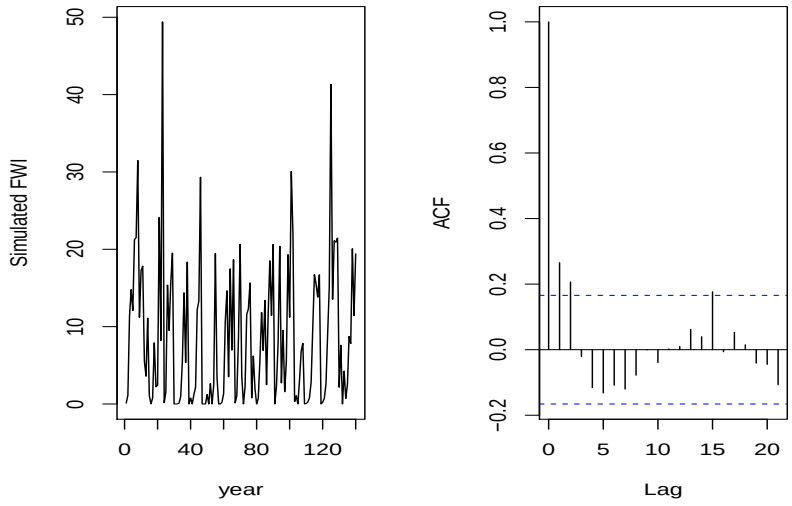

Figure 3.17: Left panel: simulated data from mixture Weibull tailed minification process using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

3.5 Random Coefficient Minification Models

We have argued on a physical basis that the tailed minification process is realistic as a model for FWI. However, if seasonality is accounted for, the first lag autocorrelation of the fitted process is not satisfactorily close to the first lag autocorrelation of the observed data.

The difficulty appears to be due to the way in which X_{n-1} enters the model by multiplication through a fixed α coefficient. In order to capture the occasional large jump in the observed series, the estimated value of α must necessarily be very large. Thus, the incidence of ε_n values being less than αX_{n-1} can be high. Because ε_n is supposed to be independent of the X_{n-1}, a large number of values of X_n will subsequently be unrelated to X_{n-1}.

In order to circumvent this problem, we will now allow α to take on values of a random variable, so that occasionally it can be large (as required by the data), but it will usually be small, in order to replicate the gradually increasing trends which are also observed in the data.

The most general form of random coefficient minification model that we will consider is

$$Y_n = \begin{cases} (\alpha_n + 1)\min(Y_{n-1}, Z_n), & \text{if } Y_{n-1} > 0. \\ W_n, & \text{if } Y_{n-1} = 0. \end{cases} \tag{3.18}$$

where α_n is a nonnegative random variable independent of Z_n, W_n and Y_{n-1}. The sequences $\{Z_n\}$, $\{W_n\}$ and $\{\alpha_n\}$ are all assumed to be i.i.d. Z_n and W_n are assumed to be tailed random variables with $p_{\mathrm{z}} = P(Z_n = 0)$, and $p_{\mathrm{w}} = P(W_n = 0)$. We denote the marginal distribution of α_n by $f(\alpha)$.

We will consider specific cases of this model for the purpose of data analysis, but beforehand, some theoretical results can be obtained for the general case.

3.5.1 Existence of a Stationary Process

Balakrishna and Jacob (2003) have proved that the minification process with positive innovations is ergodic. In this section we demonstrate the existence and uniqueness of a stationary distribution for the general random coefficient tailed minification process, using a very different proof technique.

First, note that it is possible to construct a sequence $\{Y_n\}$ such that the sequence $\{I_{Y_n=0}\}$ is stationary. To see this, note that

$$P(Y_n = 0) = P(Y_n = 0|Y_{n-1} = 0)P(Y_{n-1} = 0) + P(Y_n = 0|Y_{n-1} > 0)P(Y_{n-1} > 0))$$

or

$$P(Y_n = 0) = p_{\mathrm{W}}P(Y_{n-1} = 0) + p_{\mathrm{Z}}\left(1 - P(Y_{n-1} = 0)\right).$$

Therefore, there is a unique stationary distribution for $\{I_{Y_n=0}\}$ if and only if

$$p_{\mathrm{Y}} = p_{\mathrm{W}}p_{\mathrm{Y}} + p_{\mathrm{Z}}(1 - p_{\mathrm{Y}}) \tag{3.19}$$

has a unique solution p_{Y}, which is true when and only when

$$p_{\mathrm{W}} - p_{\mathrm{Z}} \neq 1. \tag{3.20}$$

In that case,

$$p_{\mathrm{Y}} = \frac{p_{\mathrm{Z}}}{1 + p_{\mathrm{Z}} - p_{\mathrm{W}}}.$$

Thus, the stationary distribution of $\{I_{Y_n=0}\}$ is Bernoulli with parameter p_{Y}. When $p_{\mathrm{W}} = 1$ and $p_{\mathrm{Z}} = 0$, (3.19) reduces to

$$p_{\mathrm{Y}} = p_{\mathrm{Y}}$$

which, trivially, has a solution, but it is not unique. Thus, there are infinitely many stationary distributions in this case. Note that, as a model for FWI, this is not sensible, since $p_{\mathrm{W}} = 1$ would effectively imply a constant state of flooding, while $p_{\mathrm{Z}} = 0$ would usually lead to relatively dry conditions.

For the rest of our discussion, we assume the stationarity condition (3.20), and

we set $p_{\rm Y} = P(Y_n = 0) = P(Y_{n-1} = 0)$. Then

$$\begin{aligned} P(Y_n \geq y) &= P((\alpha_n + 1)Y_{n-1} \geq y, Z_n \geq y(\alpha_n + 1)^{-1} \mid Y_{n-1} > 0)(1 - p_{\rm Y}) + \\ &\quad P(W_n \geq y)p_{\rm Y} \\ &= \int_0^\infty P(Y_{n-1} \geq y(\alpha + 1)^{-1} | Y_{n-1} > 0) P(Z_n \geq y(\alpha + 1)^{-1}) f(\alpha) d\alpha \\ &\quad (1 - p_{\rm Y}) + p_{\rm Y} P(W_n \geq y) \\ &= \int_0^\infty P(Y_{n-1} \geq y(\alpha + 1)^{-1}) P(Z_n \geq y(\alpha + 1)^{-1}) f(\alpha) d\alpha + \\ &\quad p_{\rm Y} P(W_n \geq y) \end{aligned}$$

which follows from conditioning on α_n and from the independence of Y_n and Z_n. We now set the following definitions:

$$S_{\rm Y}(y) = P(Y_n \geq y),$$

$$S_{\rm Z}(y) = P(Z_n \geq y)$$

and

$$S_{\rm W}(y) = P(W_n \geq y).$$

The above discussion suggests that $\{Y_n\}$ has a unique stationary distribution if and only if

$$S_{\rm Y}(y) = \int_0^\infty S_{\rm Y}(y(\alpha + 1)^{-1}) S_{\rm Z}(y(\alpha + 1)^{-1}) f(\alpha) d\alpha + S_{\rm W}(y) p_{\rm Y} \tag{3.21}$$

has a unique solution $S_{\rm Y}(y)$, since this would imply that $P(Y_n \geq y) = P(Y_{n-1} \geq y)$.

We will prove this in a manner analogous to proving the existence and uniqueness of the solution of a first order differential equation. To this end, let $S_0(y) =$

$p_{\mathrm{Y}} S_{\mathrm{W}}(y)$ and

$$S_n(y) = \int_0^\infty S_{n-1}(y(\alpha+1)^{-1}) S_{\mathrm{Z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha + p_{\mathrm{Y}} S_{\mathrm{W}}(y) \tag{3.22}$$

for $n = 1, 2, 3, \ldots$.

We will now demonstrate existence of a solution to (3.21) by showing that

1. $S_n(y)$ converges uniformly to $\phi(y)$, a survival function.
2. the right-hand side of equation (3.22) converges to

$$\int_0^\infty \phi(y(\alpha+1)^{-1}) S_{\mathrm{Z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha + p_{\mathrm{Y}} S_{\mathrm{W}}(y).$$

Because $S_{\mathrm{W}}(y)$ is a survival function, $S_0(y)$ must be a nonincreasing function that takes values in $[0, 1]$. Suppose $S_{n-1}(y)$ is also a nonincreasing function taking values in $[0, 1]$. We now show that this implies that $S_n(y)$ is also a survival function. Clearly, from (3.21), $S_n(y) \geq 0$, since all components of the right-hand side of that equation are nonnegative. Now, suppose that $y_2 > y_1$. Then

$S_n(y_2) - S_n(y_1) =$

$$\int_0^\infty (S_{n-1}(y_2(\alpha+1)^{-1}) S_{\mathrm{Z}}(y_2(\alpha+1)^{-1}) - S_{n-1}(y_1(\alpha+1)^{-1}) S_{\mathrm{Z}}(y_1(\alpha+1)^{-1})) f(\alpha) d\alpha$$

$$+ S_{\mathrm{W}}(y_2) p_{\mathrm{Y}} - S_{\mathrm{W}}(y_1) p_{\mathrm{Y}} \leq 0$$

because S_{n-1}, S_{Z} and S_{W} are nonincreasing. This means that $S_n(y)$ takes its maximum value on $[0, \infty)$ at $y = 0$, but

$$S_n(0) = \int_0^\infty S_{n-1}(0) S_{\mathrm{Z}}(0) f(\alpha)\, d\alpha + p_{\mathrm{Y}} S_{\mathrm{W}}(0).$$

Thus,

$$S_n(0) \leq (1 - p_{\mathrm{z}}) + (1 - p_{\mathrm{w}})p_{\mathrm{y}}$$

since $S_{n-1}(0) \leq 1$, $S_{\mathrm{z}}(0) = 1 - p_{\mathrm{z}}$, $f(\alpha)$ is a probability density function, and $S_{\mathrm{w}}(0) = 1 - p_{\mathrm{w}}$. Because

$$p_{\mathrm{y}} = \frac{p_{\mathrm{z}}}{1 - p_{\mathrm{w}} + p_{\mathrm{z}}}$$

it is clear that $S_n(0) \leq 1$. Thus, $S_n(y)$ takes values in $[0, 1]$ and is a valid survival function, and by induction, (3.22) defines a sequence of survival functions.

Furthermore, the fact that $S_n(y) \leq 1$ for all $n \geq 0$ allows us to argue that

$$\mid S_1(y) - S_0(y) \mid \leq \int_0^\infty S_0(y(\alpha+1)^{-1}) S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha \leq (1 - p_{\mathrm{z}})$$

and if

$$\mid S_n(y) - S_{n-1}(y) \mid \leq (1 - p_{\mathrm{z}})^n.$$

for $n \geq 0$, then

$$\begin{aligned} &\mid S_{n+1}(y) - S_n(y) \mid \mathrm{hspace5in} \\ &\leq \int_0^\infty \mid S_n(y(\alpha+1)^{-1}) - S_{n-1}(y(\alpha+1)^{-1}) \mid S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha \\ &\leq (1 - p_{\mathrm{z}})^n (1 - p_{\mathrm{z}}) \int_0^\infty f(\alpha)\, d\alpha = (1 - p_{\mathrm{z}})^{n+1}. \end{aligned}$$

Therefore, by induction, we have shown that

$$\mid S_n(y) - S_{n-1}(y) \mid \leq (1 - p_{\mathrm{z}})^n. \tag{3.23}$$

Our next step is to write

$$S_n(y) = p_{\mathrm{Y}} S_{\mathrm{W}}(y) + \sum_{j=1}^{n} (S_j(y) - S_{j-1}(y)).$$

The bound at (3.23) implies that

$$\sum_{j=1}^{n} \mid S_j(y) - S_{j-1}(y) \mid \leq \sum_{j=1}^{n} (1-p_{\mathrm{Z}})^j \to \frac{1-p_{\mathrm{Z}}}{p_{\mathrm{Z}}} < \infty$$

as $n \to \infty$. (Recall that $p_{\mathrm{Z}} > 0$). Therefore, there must be a function $L(y)$ for which

$$\sum_{j=1}^{n} \mid S_j(y) - S_{j-1}(y) \mid \overset{\text{uniformly}}{\longrightarrow} L(y) \leq \frac{1-p_{\mathrm{Z}}}{p_{\mathrm{Z}}}$$

as $n \to \infty$. That is,

$$\lim_{n\to\infty} S_n(y) \overset{\text{uniformly}}{\longrightarrow} p_{\mathrm{Y}} S_{\mathrm{W}}(y) + L(y).$$

We now define

$$\phi(y) = L(y) + p_{\mathrm{Y}} S_{\mathrm{W}}(y),$$

and conclude that $S_n(y)$ converges uniformly to $\phi(y)$. Because $S_n(y)$ is a survival function for each n, $\phi(y)$ must also be a survival function.

Now, we consider the right-hand side of (3.22) and show that it converges as required. First, note that

$$\mid \int_0^\infty S_n(y(\alpha+1)^{-1}) S_{\mathrm{Z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha - \int_0^\infty \phi(y(\alpha+1)^{-1}) S_{\mathrm{Z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha$$

$$\leq \int_0^\infty \mid S_n(y(\alpha+1)^{-1}) - \phi(y(\alpha+1)^{-1}) \mid S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha$$

$$\leq \sup_{\alpha\in[0,\infty)} \mid S_n(y(\alpha+1)^{-1}) - \phi(y(\alpha+1)^{-1}) \mid (1-p_{\mathrm{z}}) \int_0^\infty f(\alpha)\, d\alpha \to 0 \text{ as } n\to\infty,$$

because $f(\alpha)$ is a probability density function, $(1-p_{\mathrm{z}})$ is clearly bounded and because of our earlier result showing that $S_n(y)$ converges uniformly to $\phi(y)$.

Thus, we have established that $\phi(y)$ is a solution of equation (3.21). We will now demonstrate the uniqueness of the solution for that equation, establishing that $S_{\mathrm{Y}}(y) = \phi(y)$ is a well-defined stationary distribution. We will prove this by contradiction: suppose $\phi(y)$ and $\psi(y)$ are distinct solutions of equation (3.21). Then

$$\phi(y) - \psi(y) = \int_0^\infty \phi(y(\alpha+1)^{-1}) - \psi(y(\alpha+1)^{-1}) S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha$$

and

$$\mid \phi(y) - \psi(y) \mid \leq \int_0^\infty \mid \phi(y(\alpha+1)^{-1}) - \psi(y(\alpha+1)^{-1}) \mid S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha \quad (3.24)$$

Note that $\phi(y)$ and $\psi(y)$ are survival functions. That is, $0 \leq \phi(y) \leq 1$ and $0 \leq \psi(y) \leq 1$. Therefore we have, for all y,

$$0 \leq \mid \phi(y) - \psi(y) \mid \leq 1$$

or

$$0 \leq \mid \phi(y(\alpha+1)^{-1}) - \psi(y(\alpha+1)^{-1}) \mid \leq 1$$

and, because of (3.24),

$$0 \leq \mid \phi(y) - \psi(y) \mid \leq \int_0^\infty S_{\mathrm{z}}(y(\alpha+1)^{-1}) f(\alpha)\, d\alpha \leq (1-p_{\mathrm{z}}) \int_0^\infty f(\alpha)\, d\alpha = 1 - p_{\mathrm{z}}.$$

That is, for any $\alpha > 0$,

$$| \phi(y(\alpha+1)^{-1}) - \psi(y(\alpha+1)^{-1}) | \leq 1 - p_{\mathrm{z}}. \tag{3.25}$$

Using result (3.25) in equation (3.24), we see that

$$| \phi(y) - \psi(y) | \leq (1 - p_{\mathrm{z}})^2$$

Repeating this process, inductively, we have

$$| \phi(y) - \psi(y) | \leq (1 - p_{\mathrm{z}})^n.$$

Observe that

$$\lim_{n\to\infty} (1 - p_{\mathrm{z}})^n = 0.$$

Therefore, we have

$$\phi(y) = \psi(y).$$

Thus, equation (3.21) possesses a unique solution. The random coefficient tailed minification process has a unique stationary distribution.

3.5.2 The Weibull Tailed Mixture Minification Process

For a positive power P, consider

$$Y_n = X_n^{\mathrm{P}} \tag{3.26}$$

where

$$X_n = \begin{cases} (\alpha_n + 1)\min(X_{n-1}, \alpha_n^{-1}\varepsilon_n), & \text{if } X_{n-1} > 0. \\ \delta_n, & \text{if } X_{n-1} = 0. \end{cases} \tag{3.27}$$

Again, α_n is a nonnegative random variable independent of ε_n, δ_n and X_{n-1}. The sequences $\{\varepsilon_n\}$ and $\{\delta_n\}$ are assumed to be i.i.d. exponential tailed random variables with parameters λ and γ, and $p_\varepsilon = P(\varepsilon_n = 0)$ and $p_\delta = P(\delta_n = 0)$: $\mathrm{ET}(\lambda, p_\varepsilon)$ and $\mathrm{ET}(\gamma, p_\delta)$. The marginal distribution of α_n is $f(\alpha)$.

3.5.3 Stationarity

We have already considered the stationarity question in general, but we can say more in this special case.

If $\gamma = \lambda$, it is possible to define a stationary process via (3.27), for any pdf $f(\alpha)$ having support on the nonnegative real numbers. To see this, suppose that

$$P(X_{n-1} \geq x) = (1 - p_X)e^{-\lambda x},$$

where $p_X = p_\varepsilon(1 + p_\varepsilon - p_\delta)^{-1}$. It follows that

$$\begin{aligned} P(X_n \geq x) &= p_X(1 - p_\delta)e^{-\lambda x} + \\ &\quad \int_0^\infty P\left(X_{n-1} \geq x(\alpha+1)^{-1}, \varepsilon_n \geq x\lambda(\alpha+1)^{-1} | X_{n-1} > 0\right) f(\alpha)\, d\alpha \\ &= p_X(1 - p_\delta)e^{-\lambda x} + (1 - p_X)(1 - p_\varepsilon)e^{-\lambda x} \int_0^\infty f(\alpha)\, d\alpha \\ &= (1 - p_X)e^{-\lambda x}. \end{aligned}$$

Therefore, $\{X_n\}$ is a stationary exponential tailed process. Furthermore, when $Y_n = X_n^{\mathrm{P}}$, $\{Y_n\}$ is a stationary Weibull tailed process with parameters p_{Y}, λ and P: $\mathrm{WT}(\lambda, p_{\mathrm{Y}}, \mathrm{P})$.

When $\gamma \neq \lambda$ in (3.27), and stationarity holds, the survival probability $P(X_n \geq x)$ must satisfy the following equation:

$$P(X_n \geq x) = (1 - p_\delta)p_X e^{-\gamma x} + (1 - p_\varepsilon) \int_0^\infty P\left(X_n \geq x(\alpha+1)^{-1}\right) e^{-\frac{\alpha\lambda x}{\alpha+1}} f(\alpha) d\alpha. \tag{3.28}$$

The steps to derive (3.28) are as follows:

$$\begin{aligned}
P(X_n \geq x) &= P(X_n \geq x | X_{n-1} = 0) P(X_{n-1} = 0) + \\
&\quad P(X_n \geq x) | X_{n-1} > 0) P(X_{n-1} > 0) \\
&= p_X (1 - p_\delta) e^{-\gamma x} + \\
&\quad (1 - p_X) P(X_{n-1} \geq x(\alpha_n + 1)^{-1}, \varepsilon_n \geq \alpha_n x (\alpha_n + 1)^{-1} | X_{n-1} > 0) \\
&= \int_0^\infty P(X_{n-1} \geq x(\alpha + 1)^{-1} | X_{n-1} > 0)(1 - p_\varepsilon) e^{-\frac{\alpha \lambda x}{\alpha + 1}} f(\alpha) d\alpha \\
&\quad (1 - p_X) + p_X (1 - p_\delta) e^{-\gamma x} \\
&= p_X (1 - p_\delta) e^{-\gamma x} + \\
&\quad (1 - p_\varepsilon) \int_0^\infty P(X_{n-1} \geq x(\alpha + 1)^{-1}) e^{-\frac{\alpha \lambda x}{\alpha + 1}} f(\alpha) d\alpha
\end{aligned}$$

If the process is stationary, then

$$P(X_{n-1} \geq x(\alpha + 1)^{-1}) = P(X_n \geq x(\alpha + 1)^{-1}).$$

Thus, we obtain (3.28).

In the Weibull tailed case, i.e. $Y_n = X_n^{1/\mathrm{P}}$, have

$$\begin{aligned}
P((Y_n \geq y) &= P(X_n \geq y^{1/\mathrm{P}}) \\
&= (1 - p_\delta) p_X e^{-\gamma y^{1/\mathrm{P}}} + \qquad (3.29) \\
&\quad (1 - p_\varepsilon) \int_0^\infty P\left(X_n \geq y^{1/\mathrm{P}} (\alpha + 1)^{-1}\right) e^{-\frac{\alpha \lambda y^{1/\mathrm{P}}}{\alpha + 1}} f(\alpha) d\alpha. \qquad (3.30)
\end{aligned}$$

so that

$$P((Y_n \geq y) = (1-p_\delta)p_Y e^{-\gamma y^{1/\mathrm{P}}} + (1-p_\varepsilon)\int_0^\infty P\left(Y_n \geq y(\alpha+1)^{-\mathrm{P}}\right) e^{-\frac{\alpha\lambda y^{1/\mathrm{P}}}{\alpha+1}} f(\alpha)d\alpha. \quad (3.31)$$

3.5.4 Conditional Expectation

One step ahead predictions are often made, using the conditional expectation of X_n, given X_{n-1}. This expectation can also be used for model-checking or for least-squares parameter estimation.

For the exponential case, we have

$$E[X_n|X_{n-1}=x] = \begin{cases} E[(\alpha_n+1)xI_{\{\varepsilon_n \geq \alpha_n x\}} + (\alpha_n+1)\alpha_n^{-1}\varepsilon_n I_{\{\varepsilon_n < \alpha_n x\}}], & \text{if } x>0. \\ (1-p_\delta)\gamma^{-1}, & \text{if } x=0. \end{cases}$$

When $x > 0$,

$$E[X_n|X_{n-1}=x] = \int_0^\infty (\alpha+1)xe^{-\lambda\alpha x} f(\alpha)\, d\alpha(1-p_\varepsilon)+$$

$$(1-p_\varepsilon)\int_0^\infty (\alpha+1)(\lambda\alpha)^{-1}(1-\lambda\alpha x e^{-\lambda\alpha x} - e^{-\lambda\alpha x}) f(\alpha)\, d\alpha$$

so that

$$E[X_n|X_{n-1}=x] = (1-p_\varepsilon)\int_0^\infty (\alpha+1)(\alpha\lambda)^{-1}(1-e^{-\lambda\alpha x}) f(\alpha)\, d\alpha \quad (3.32)$$

It should be noted that when $f(\alpha)$ is an exponential probability density function, this expression simplifies substantially. However, our main application assumes a lognormal form for $f(\alpha)$, so further simplication is not possible.

For the Weibull case, i.e. $Y_n = X_n^{\mathrm{P}}$, we have

$$E[Y_n|Y_{n-1}=0] = E[\delta^{\mathrm{P}}] = (1-p_\delta)E[T^{\mathrm{P}}]$$

where T is exponential with parameter γ. Therefore,

$$E[Y_n|Y_{n-1}=0] = (1-p_\delta)\Gamma(\mathrm{P}+1)\gamma^{-\mathrm{P}}.$$

When $x > 0$,

$$\begin{aligned} E[Y_n|Y_{n-1}=x] &= E[X_n^{\mathrm{P}}|X_{n-1}=x^{1/\mathrm{P}}] \\ &= \int_0^\infty (\alpha+1)^{\mathrm{P}} x^{\mathrm{P}} e^{-\lambda\alpha x} f(\alpha)\, d\alpha (1-p_\varepsilon) \\ &\quad + \int_0^\infty (\alpha+1)^{\mathrm{P}} \alpha^{-\mathrm{P}} \lambda^{-1} \left(\int_{\alpha x}^\infty \varepsilon^{\mathrm{P}} e^{-\lambda\varepsilon} d(\varepsilon) \right) f(\alpha)\, d\alpha (1-p_\varepsilon) \end{aligned} \tag{3.33}$$

3.5.5 Likelihood

The process $\{Y_n\}$ following (3.26) and (3.27) is Markovian. Therefore, the likelihood for a sample $y_1, y_2, \ldots, y_n$ is

$$L(y_1, y_2, \ldots, y_n) = \prod_{i=2}^{n} f_{Y_i|Y_{i-1}}(y_i|y_{i-1}) \cdot f_{Y_1}(y_1). \tag{3.34}$$

The corresponding log likelihood is then given by

$$\log L = \sum_{i=1}^{n} \log\left(f_{Y_i|Y_{i-1}}(y_i|y_{i-1}) \right) + \log\left(f_{Y_1}(y_1) \right). \tag{3.35}$$

Thus, in order to calculate the log likelihood for given data, it is necessary to obtain an expression for the conditional density of Y_n, given Y_{n-1} as well as the marginal

density of Y_1.

The marginal density for Y_1 is obtained by differentiating the survival function $P(Y_1 \geq y_1)$ (obtained in the previous section under the condition of stationarity) with respect to y_1 and multiplying through by -1. When $\lambda \neq \gamma$, a numerical technique is needed to calculate the derivative.

We obtain the needed conditional densities in the next subsections for the exponential tailed and Weibull tailed cases.

3.5.5.1 Exponential Tailed Mixture Case

We assume first that P can be specified beforehand. Then $X_n = Y_n^{1/\mathrm{P}}$ follows a mixture of exponential tailed distributions, and we can condition on α_n to show that

$$P(X_n \geq y | X_{n-1} = x) = \int_0^\infty P(X_{n-1} \geq y(\alpha+1)^{-1}) f(\alpha)\, d\alpha$$

when $x > 0$ and $y > 0$,. The probability statement in the integrand is equivalent to

$$P(\varepsilon_n \geq y\alpha(\alpha+1)^{-1}) I(x \geq y(\alpha+1)^{-1})$$

since ε_n and X_{n-1} are independent. Note that $x \geq y(\alpha+1)^{-1}$ if and only if $\alpha \geq \frac{y}{x} - 1$. Therefore, we have

$$P(X_n \geq y | X_{n-1} = x) = \int_{\max(0, y/x-1)}^\infty P(\varepsilon_n \geq y\lambda(\alpha+1)^{-1}) f(\alpha)\, d\alpha.$$

Note that it is possible that $y < x$.

The exponential tailed distribution assumption on ε_n leads to

$$P(X_n \geq y | X_{n-1} = x) = \int_{\max(0, y/x-1)}^\infty (1 - p_\varepsilon) e^{-\lambda y \alpha (\alpha+1)^{-1}} f(\alpha)\, d\alpha \qquad (3.36)$$

when $x > 0$ and $y > 0$.

For the case where $x = 0$, it is easy to see that

$$P(X_n \geq y | X_{n-1} = 0) = \begin{cases} (1 - p_\delta)e^{-\gamma y}, & \text{if } y > 0. \\ 1, & \text{if } y = 0. \end{cases}$$

Finally,

$$P(X_n \geq y | X_{n-1} = 0) = 1.$$

To obtain the conditional pdf of X_n, given $X_{n-1} = x$ from the above conditional survival functions, we must differentiate with respect to y and multiply through by -1, noting the point masses at 0 which imply that

$$f_{X_n|X_{n-1}}(y|x) = p_\delta$$

when $x = y = 0$, and $f_{X_n|X_{n-1}}(y|x) = p_\varepsilon$ when $x > 0$ and $y = 0$.

When $y > 0$ and $x = 0$, we easily obtain

$$f_{X_n|X_{n-1}}(y|x) = (1 - p_\delta)\gamma e^{-\gamma y}.$$

The case when $y > 0$ and $x > 0$ is a little more complicated. In that case,

$$\begin{aligned} f_{X_n|X_{n-1}}(y|x) \quad = \quad & \int_{\max(0, \frac{y}{x} - 1)}^{\infty} (1 - p_\varepsilon)\alpha(\alpha + 1)^{-1}\lambda e^{-\frac{\alpha\lambda y}{\alpha+1}} f(\alpha)\, d\alpha + \\ & \mathrm{I}(1 - p_\varepsilon)e^{-\frac{\alpha'\lambda y}{\alpha'+1}} f(\alpha')/x \end{aligned}$$

where $\alpha' = y/x - 1$ and I is 1, if $y \geq x$ and 0, otherwise.

This conditional density can be inserted into the likelihood or log likelihood expressions given at (3.34) or (3.35). For example, the log likelihood given data of the form $y_1, y_2, \ldots, y_n$ and known Weibull parameter P is

$$\log L = \sum_{i=1}^{n} \log\left(f_{X_i|X_{i-1}}(y_i^{1/\mathrm{P}} | y_{i-1}^{1/\mathrm{P}})\right) + \log\left(f_{X_1}(y_1^{1/\mathrm{P}})\right). \tag{3.37}$$

3.5.5.2 Weibull Tailed Mixture Case

Now, we consider the more general case of the model specified by (3.26) and (3.27) where P is unknown. Since $Y_n = X_n^{\mathrm{P}}$, our earlier reasoning leads to

$$\begin{aligned} P(Y_n \geq y | Y_{n-1} = x) &= P(X_n \geq y^{1/\mathrm{P}} | X_{n-1} = x^{1/\mathrm{P}}) \\ &= (1 - p_\varepsilon) \int\limits_{\max(0,(\frac{y}{x})^{1/\mathrm{P}}-1)}^{\infty} e^{-\frac{\alpha\lambda y^{1/\mathrm{P}}}{\alpha+1}} f(\alpha)\, d\alpha \end{aligned} \quad (3.38)$$

when $x > 0$ and $y > 0$.

$$P(Y_n > y | Y_{n-1} = x) = (1 - p_\delta)\gamma e^{-\gamma y^{1/\mathrm{P}}}$$

when $y > 0$ and $x = 0$ and $P(Y_n > y | Y_{n-1} = x) = 1$, when $y = 0$, $x \geq 0$. The conditional pdf is then

$$f_{Y_n|Y_{n-1}}(y|x) = \int\limits_{\max(0,(y/x)^{1/\mathrm{P}}-1)}^{\infty} \mathrm{P}^{-1}(1 - p_\varepsilon)(\alpha(\alpha+1)^{-1})\lambda y^{1/\mathrm{P}-1} e^{-\frac{\alpha\lambda y^{1/\mathrm{P}}}{\alpha+1}} f(\alpha)\, d\alpha$$

$$\mathrm{I}(1 - p_\varepsilon) e^{-\frac{\alpha'\lambda y^{1/\mathrm{P}}}{\alpha'+1}} f(\alpha')(Px)^{-1}(y/x)^{1/\mathrm{P}-1})$$

for $x > 0$, $y > 0$, where $\mathrm{I} = 1$, when $y/x > 1$ and 0, otherwise and $\alpha' = (y/x)^{1/\mathrm{P}-1}$. When $y > 0$ and $x = 0$,

$$f_{Y_n|Y_{n-1}}(y|x) = \gamma(\mathrm{P})^{-1} y^{1/\mathrm{P}-1} e^{-\gamma y^{1/\mathrm{P}}}$$

and

$$f_{Y_n|Y_{n-1}}(0|x) = p_\varepsilon$$

when $x > 0$ and

$$f_{Y_n|Y_{n-1}}(0|0) = p_\delta.$$

This conditional density can be inserted into the likelihood or log likelihood expressions given at (3.34) or (3.35).

3.5.6 Estimating the Process Parameters

The log likelihood can be maximized numerically to give estimates of p_δ, p_ε, λ, γ and P (in the Weibull case), as well as the parameters of the pdf $f(\alpha)$.

The distribution of α must be pre-specified. Since α must be nonnegative, some possible models are exponential, gamma, Weibull, lognormal. Models such as a scaled beta distribution could be considered, but this imposes an upper bound on α which is not necessarily realistic. We have considered two cases of $f(\alpha)$: exponential (with rate parameter β), and lognormal (with parameters μ and σ, the mean and standard deviation on the log scale). Since the lognormal distribution gives a better fit to the data, we restrict our discussion to this case. We should also note that taking $\mu = 0$ appears to work better than to allow μ to be estimated. This may be due to an identifiability problem which should be addressed in a future study.

3.5.6.1 Application to the 1987 Dryden Data

Parameter estimates for the random coefficient exponential tailed mixture minification model fit to the 1987 Dryden data are as follows: $\widehat{p_\varepsilon} = 0.149$, $\widehat{\lambda} = 0.103$, $\widehat{\gamma} = 0.607$, $\widehat{p_\delta} = 0.32$ and $\widehat{\sigma} = 1.36$. Figure 3.18 shows a typical run of data simulated from this fitted model, together with the corresponding autocorrelation function plot. The FWI values are fairly realistic, but the autocorrelations at small lags do not match the autocorrelations in the observed data. This model is not satisfactory for these data.

Parameter estimates for the random coefficient Weibull tailed mixture minification model fit to the 1987 Dryden data are as follows, when the power P is fixed at 0.833: $\widehat{p_\varepsilon} = 0.149$, $\widehat{\alpha} = 8$, $\widehat{\lambda} = 0.057$, $\widehat{\gamma} = 0.527$, $\widehat{p_\delta} = 0.32$ and $\widehat{\sigma} = 1.481$. Figure 3.19 shows a typical run of data simulated from this fitted model, together with the corresponding autocorrelation function plot. The FWI values are fairly realistic, if somewhat low, while the autocorrelations at small lags match the autocorrelations in the observed data adequately. This model would be viewed as acceptable by this criterion and will be subject to further scrutiny subsequently.

It is also possible to estimate the Weibull power parameter. Parameter estimates for the random coefficient Weibull tailed mixture minification model fit to the 1987 Dryden data are as follows, when the power P is estimated: $\mathrm{P} = 0.886, p_\varepsilon = 0.149, p_\delta = 0.32, \lambda = 0.0705, \gamma = 0.555$, and $\sigma = 1.42$. Figure 3.20 shows a typical run of data simulated from this fitted model, together with the corresponding autocorrelation function plot. Again, the FWI values are quite realistic, while the

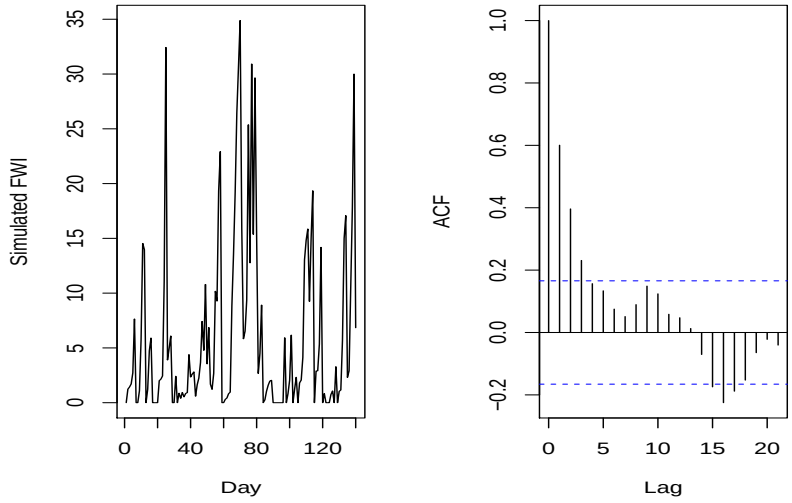

Figure 3.18: Left panel: simulated run of random coefficient exponential tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

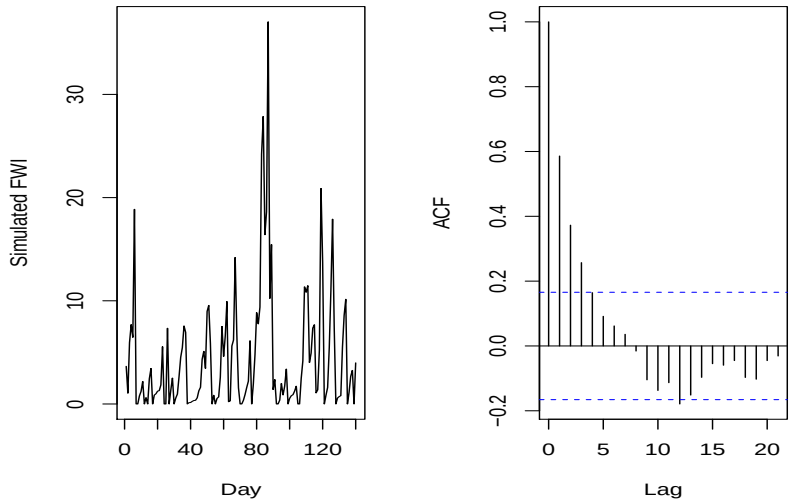

Figure 3.19: Left panel: simulated run of random coefficient Weibull tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season, with P = 5/6. Right panel: autocorrelation estimates for the simulated data.

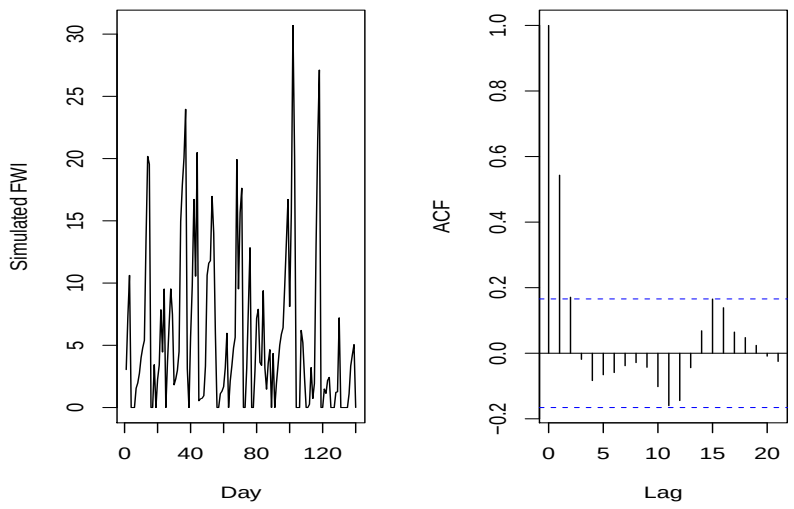

Figure 3.20: Left panel: simulated run of random coefficient Weibull tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 1987 fire season. In this case, the power parameter P has been estimated. Right panel: autocorrelation estimates for the simulated data.

autocorrelations at small lags match the autocorrelations in the observed data adequately.

3.5.6.2 Accounting for Seasonality

Based on our earlier observations as well as those of Martell (1999), we have reason to believe that seasonality plays a role, even with the random coefficient model we are considering.

Table 3.6 contains maximum likelihood estimates of $p_\varepsilon, \lambda, p_\delta, \gamma$ and σ for each of Martell's subseasons for the 1987 Dryden FWI data. The innovations ε_n and δ_n are assumed to be exponential here. The estimates of the first four parameters behave similarly to the corresponding estimates for other seasonal models we have considered. The estimate of σ is interesting; it remains constant through the first two subseasons before dropping substantially in the third.

Tables 3.7 and 3.8 contain the parameter estimates for the seasonal Weibull model. In the former case, P has been set to 5/6 while in the latter case, P has been estimated by maximum likelihood. When P was taken to be constant over the whole season, the estimates of σ have a similar pattern to the exponential case. When P is

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.09	0.09	0.12
p_δ	0.29	0.36	0.20
γ	0.67	0.67	0.67
σ	1.36	1.38	0.65

Table 3.6: Parameter estimates for the random coefficient exponential tailed mixture model fit to each of the 3 subseasons of the 1987 Dryden FWI data.

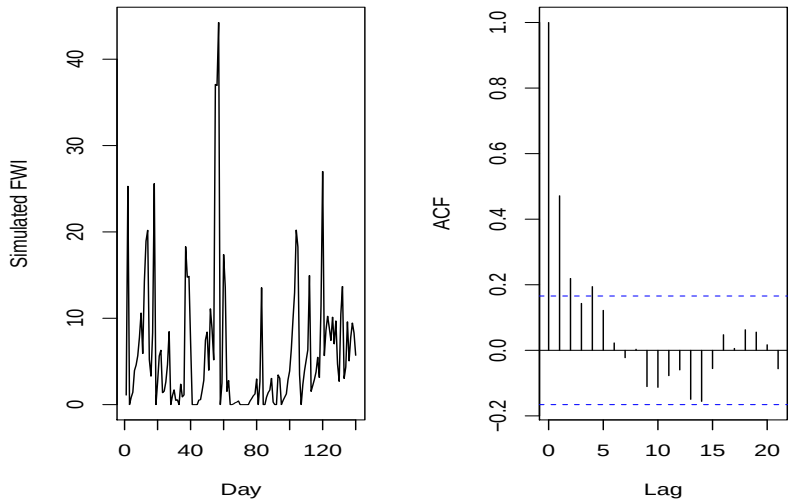

Figure 3.21: Left panel: simulated run of random coefficient exponential tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. Right panel: autocorrelation estimates for the simulated data.

allowed to vary from subseason to subseason, it has a value near 1 for the first two subseasons and then a sharp drop in the third subseason. In this case the estimates of σ are somewhat more stable.

Figures 3.21, 3.22 and 3.23 contain the simulated trace and autocorrelation plots for the three fitted seasonal models we have considered.

All three trace plots seen to be realistic, displaying patterns of 0's and runs of nonzero values that could be seen in real data. Figure 3.22 contains a relatively large simulated FWI value, indicating a possible hint of trouble. The autocorrelation plots match the observed autocorrelation plots very well, especially in the Weibull cases.

As a model for the 1987 Dryden FWI data, our informal diagnostic checks

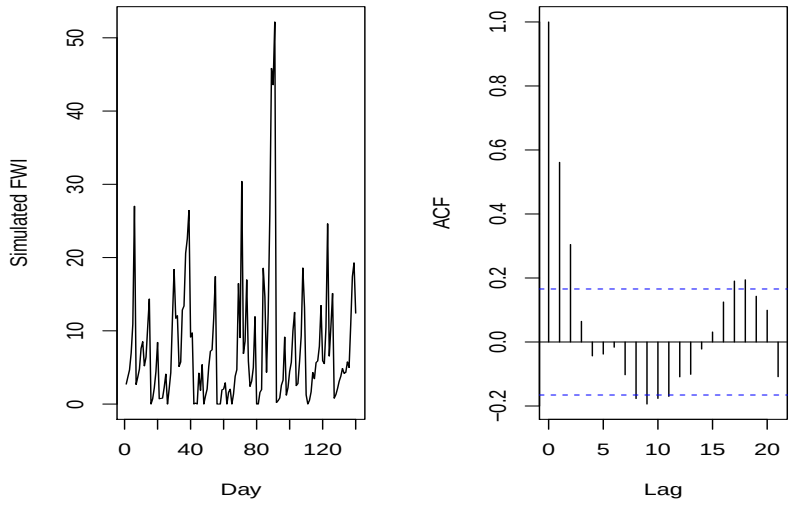

Figure 3.22: Left panel: simulated run of random coefficient Weibull tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. The Weibull parameter has been fixed at 5/6. Right panel: autocorrelation estimates for the simulated data.

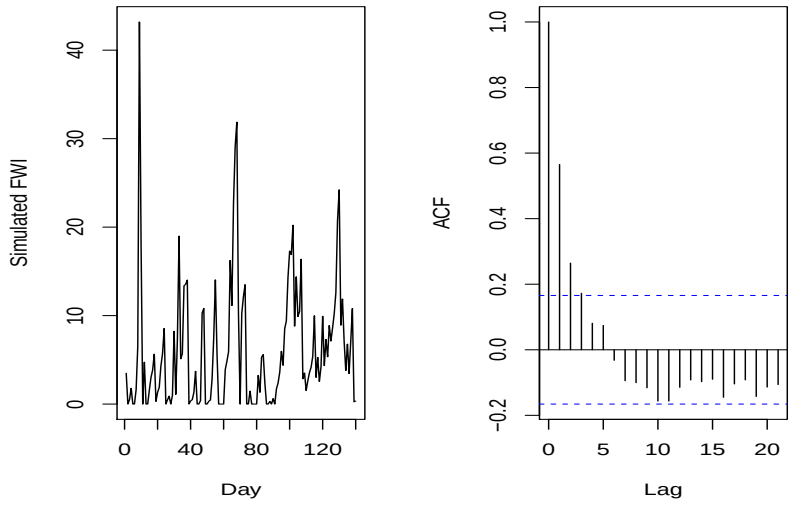

Figure 3.23: Left panel: simulated run of random coefficient Weibull tailed mixture minification model using parameters estimated from FWI observations at Dryden, Ontario, for the 3 subseasons of the 1987 fire season. The Weibull parameter has been estimated using Least-Squares. Right panel: autocorrelation estimates for the simulated data.

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.05	0.05	0.07
p_δ	0.29	0.36	0.20
γ	0.59	0.59	0.59
σ	1.57	1.49	0.83

Table 3.7: Parameter estimates for the random coefficient Weibull tailed mixture model fit to each of the 3 subseasons of the 1987 Dryden FWI data, assuming P = 5/6

	Subseason 1	Subseason 2	Subseason 3
p_ε	0.11	0.16	0.07
λ	0.11	0.09	0.02
p_δ	0.29	0.36	0.20
γ	0.69	0.69	0.69
σ	1.30	1.38	1.17
P	1.07	0.99	0.57

Table 3.8: Parameter estimates for the random coefficient Weibull tailed mixture model fit to each of the 3 subseasons of the 1987 Dryden FWI data. In this case, P has been estimated by maximum likelihood.

indicate that the seasonal random coefficient Weibull tailed minification model may be a suitable candidate.

In the next section we introduce additional diagnostic checks and consider the larger set of data covering about 40 fire seasons at six Ontario weather stations.

3.6 Model Checking Using the Larger FWI Data Set

The seasonal random coefficient Weibull tailed minification process appears to be a candidate model for the 1987 Dryden FWI time series. In this section, we will consider the seasonal and nonseasonal versions of this model as they apply to other years and other weather stations in Ontario. We also consider additional techniques for model checking, since it is well known that assessing linear dependence only is insufficient in nonlinear models.

3.6.1 Checks Based on Autocorrelations

We will first check lag 1 autocorrelations on all years of data and at each weather station at our disposal. For each of the 6 weather stations, the random coefficient

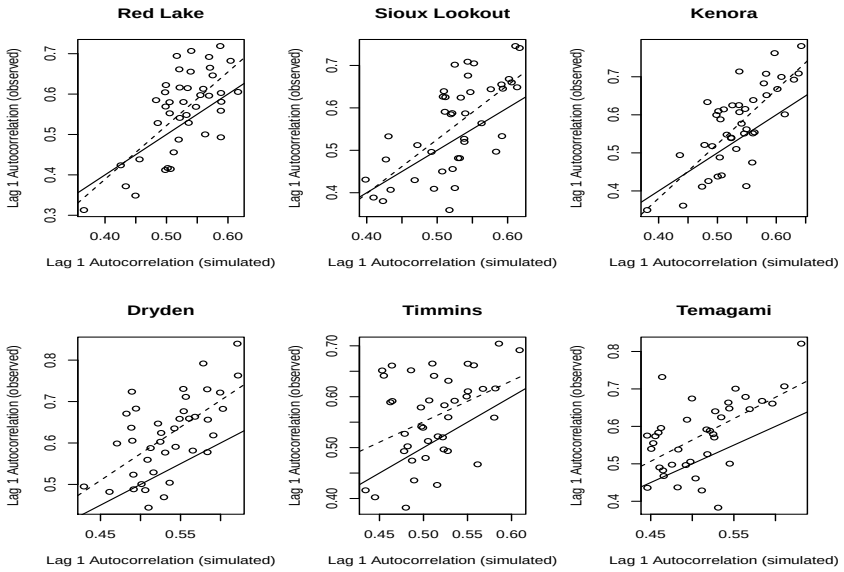

Figure 3.24: Lag 1 autocorrelations for each year of observed FWI data at each of 6 Ontario weather stations plotted against the corresponding model-based lag 1 autocorrelations. The 45° reference line is shown as a solid curve and the least-squares fitted line is shown as a dashed curve.

minification model has been fit to each year of FWI data. Using the resulting fitted models, we have simulated 10000 observations in order to calculate approximations to the lag 1 autocorrelations. Lag 1 autocorrelations have also been estimated for each year of observed FWI data.

Figure 3.24 displays a plot of the estimated lag 1 autocorrelations versus the corresponding model-based estimates, for each year of data, at each of the 6 weather stations. If the first lag autocorrelations in the observed data are being modelled adequately by the random coefficient minification model, we would expect the plotted points to scatter along the 45° line, which has been plotted as a solid line. The least-squares fitted line has also been plotted to aid in making the comparison.

As a further aid, we have also constructed a similar plot using simulated datasets which emulate the original FWI datasets in Figure 3.25. The simulated datasets were generated from the fitted models, with the annual sample sizes taken to match those for the original series. Again, lag 1 autocorrelations have been estimated for the simulated FWI datasets and plotted against the lag 1 autocorrelations calculated from sequences of 10000 simulated observations. In this case, we are making a comparison where we know that there should be an exact match, because the models being compared are identical.

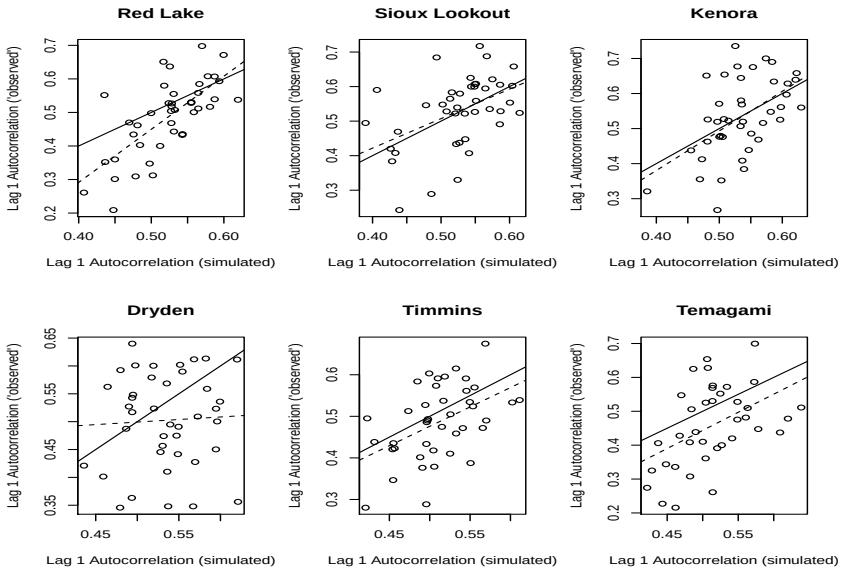

Figure 3.25: Lag 1 autocorrelations for each year of simulated FWI data as if it were taken at each of 6 Ontario weather stations plotted against the corresponding model-based lag 1 autocorrelations using simulated sets of 10000 observations. The 45° reference line is shown as a solid curve and the least-squares fitted line is shown as a dashed curve.

The second set of plots suggests how much vertical scatter should be expected on the first set of plots. Because the simulation sample sizes are very large, little horizontal scatter is expected. The plots in Figure 3.25 indicate how close the least-squares and 45° lines should be in Figure 3.25, if the random coefficient minification model is adequate for the data.

With the exception of the Dryden data, it appears that the lag 1 autocorrelation behaviour is captured reasonably well by the fitted models, although there is some evidence that the autocorrelations are more variable than expected if the models are true. The nonseasonal minification model is clearly underestimating the autocorrelation at Dryden.

Figures 3.26 and 3.27 contain the analogous plots for the case where Martell's subseasons have been taken into account in the random coefficient minification models.

There are some slight differences in how these plots have been constructed that should be mentioned. In 3.26 we are still plotting the estimated lag 1 autocorrelations from the observed FWI data for each year and weather station, these autocorrelations are plotted against autocorrelations calculated for data simulated from the seasonal models. These latter datasets are constructed somewhat differently in order for the

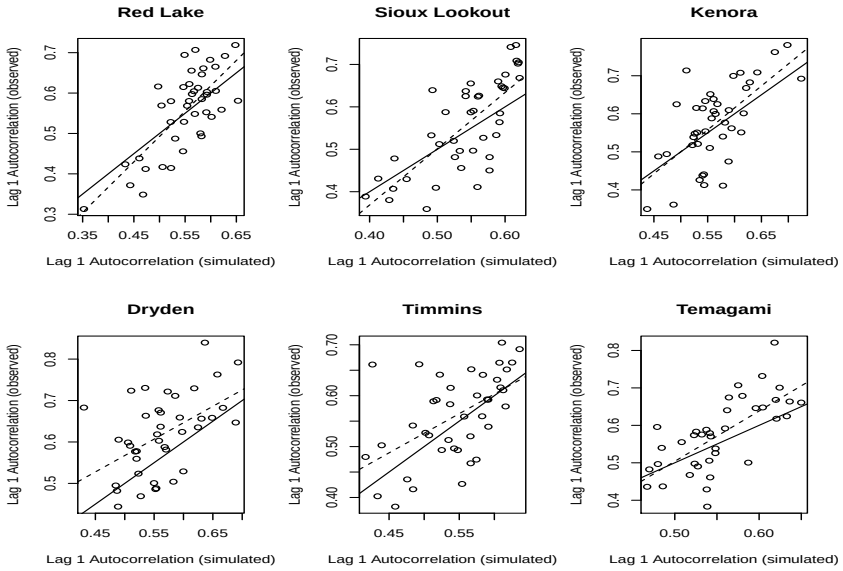

Figure 3.26: Lag 1 autocorrelations for each year of observed FWI data at each of 6 Ontario weather stations plotted against the corresponding (seasonal) model-based lag 1 autocorrelations. The 45° reference line is shown as a solid curve and the least-squares fitted line is shown as a dashed curve.

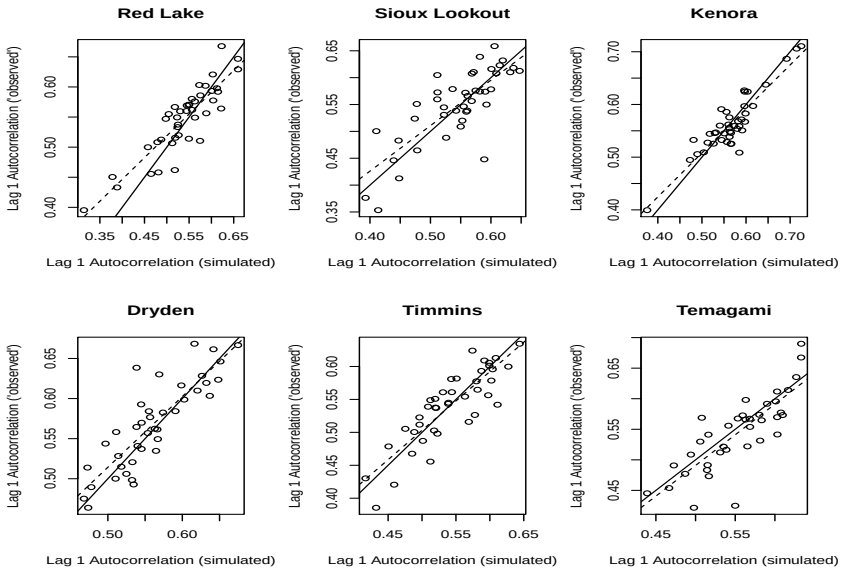

Figure 3.27: Lag 1 autocorrelations for each year of simulated seasonal FWI data as if it were taken at each of 6 Ontario weather stations plotted against the corresponding (seasonal) model-based lag 1 autocorrelations using simulated sets of 10000 observations. The 45° reference line is shown as a solid curve and the least-squares fitted line is shown as a dashed curve.

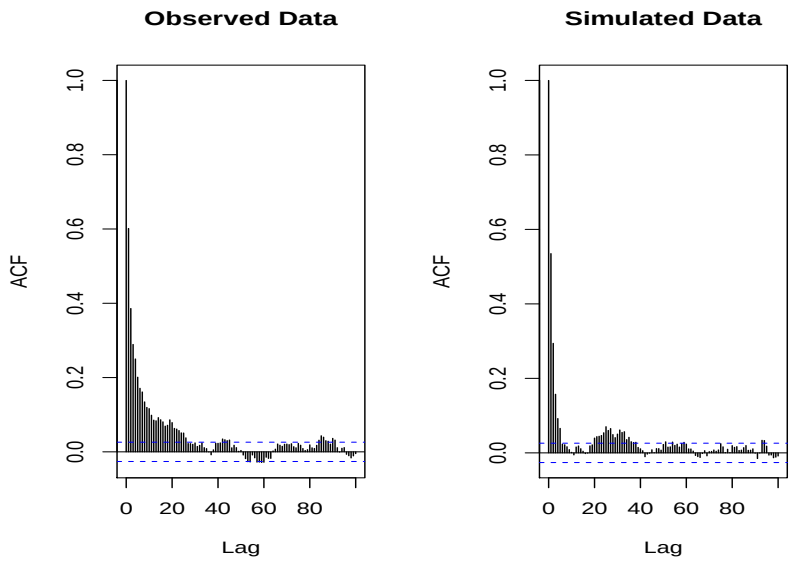

Figure 3.28: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Red Lake. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Red Lake FWI series.

seasonal effect on the autocorrelations to be demonstrated. Specifically, for each year, 20 "years" of FWI data have been simulated from the fitted seasonal model; the lag 1 autocorrelation has been calculated from this simulated series.

Now there is a closer match between the sets of plots based on comparisons of observed and expected autocorrelations and sets of plots based on comparisons of simulated and expected autocorrelations, even in the case of the Dryden data. Thus, the seasonal minification models offer clear improvement over the nonseasonal models.

Figures 3.28 through 3.33 contain long run autocorrelation plots for all years of FWI data from the six weather stations. In the left panel of each figure are plotted the autocorrelations at lags 1 through 100 for the entire set of observed data. In the right panel, the corresponding autocorrelations are plotted for data simulated from models fitted to each year of data. Seasonality is not assumed here.

In several of the figures, the autocorrelations at lags 5 through 20 for the simulated data tend to be much less than the corresponding autocorrelations for the observed data. The implication is that the model is not adequate for the data.

Figures 3.34 through 3.39 contain the analogous long run autocorrelation plots for all years of FWI data from the six weather stations using the seasonal random coefficient minification model.

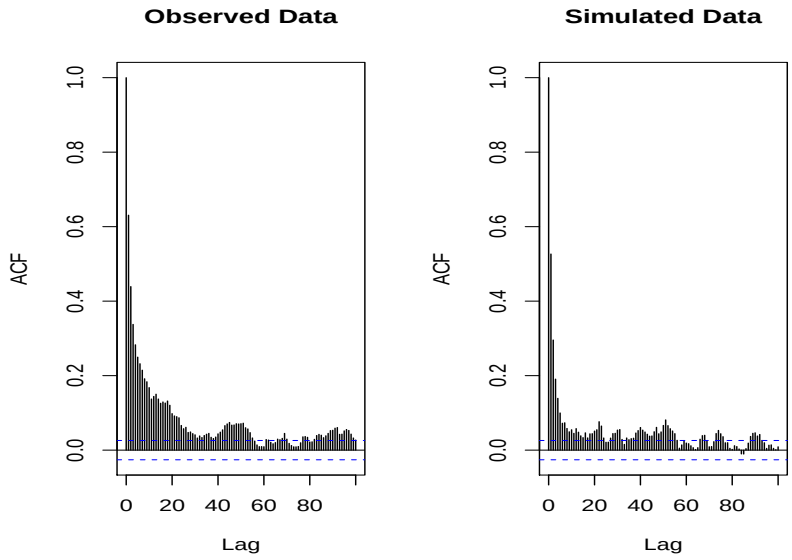

Figure 3.29: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Sioux Lookout. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Sioux Lookout FWI series.

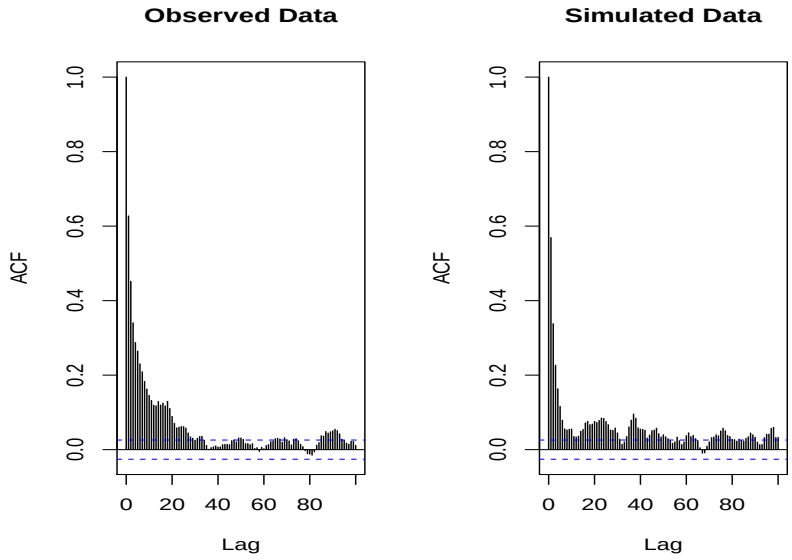

Figure 3.30: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Kenora. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Kenora FWI series.

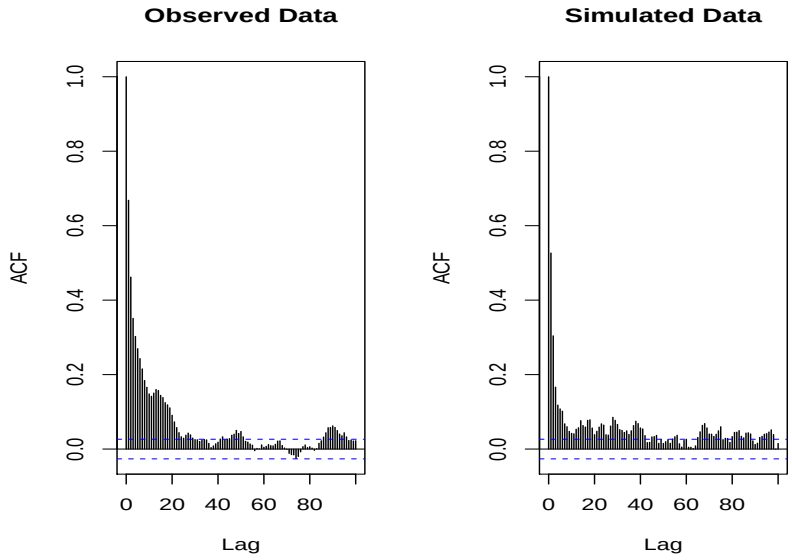

Figure 3.31: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Dryden. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Dryden FWI series.

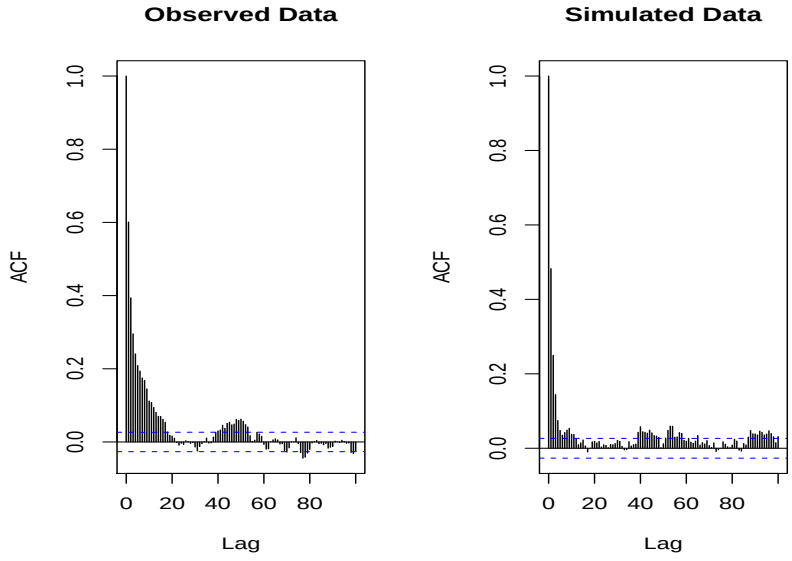

Figure 3.32: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Timmins. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Timmins FWI series.

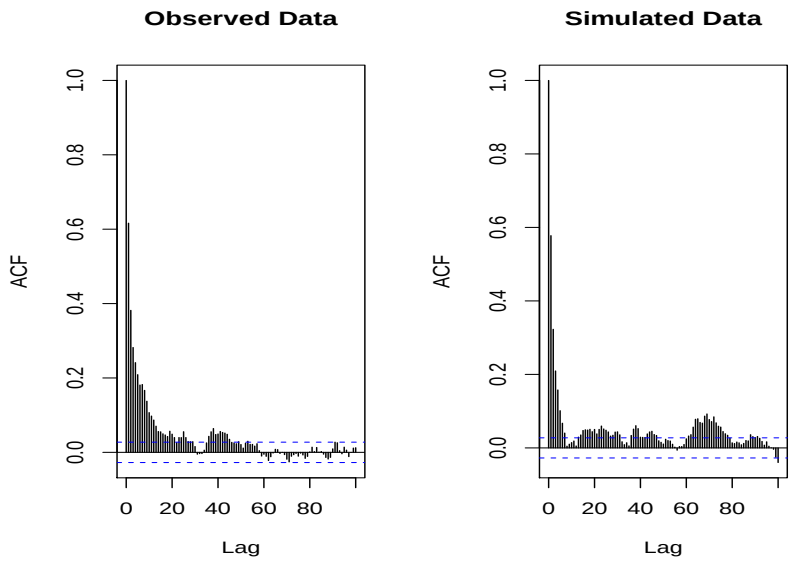

Figure 3.33: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Temagami. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of models fitted to the Temagami FWI series.

The simulated and observed autocorrelation plots match fairly closely at lags up to approximately 30. The implication is that seasonality corrects the model at higher lags. The seasonal random coefficient minification process accurately models the linear components of the dependence in the observed data.

3.6.2 Checks Based on Residuals

When possible, time series and regression models are checked against data using a variety of residual plots. In this section, we apply two graphical residual diagnostic plots as checks on the appropriateness of the random coefficient minification process as a model for the Ontario FWI time series.

The first graphic is a type of one-step-ahead prediction residuals, based on conditional expectations.

For the exponential case, we compute fitted values using the one-step-ahead conditional expectations, evaluated at their maximum likelihood estimates. Residuals are calculated as

$$R_n = X_n \div \hat{E}[X_n|X_{n-1}]$$

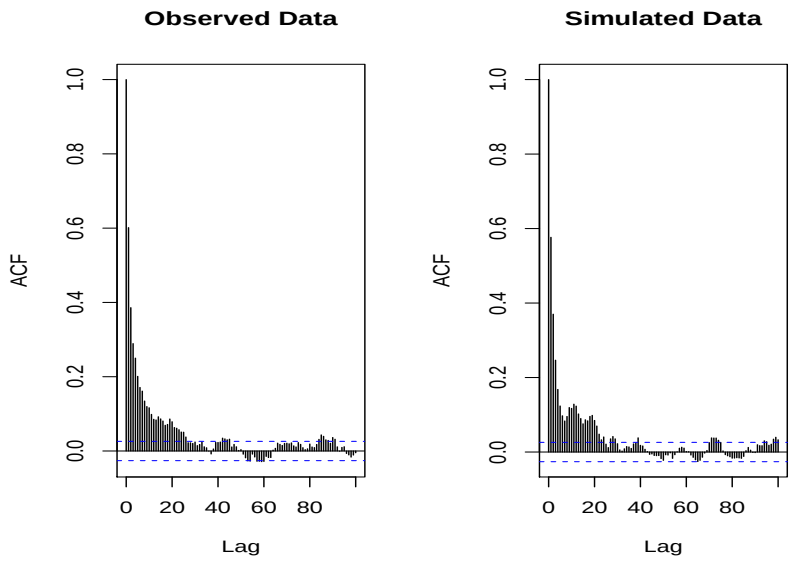

Figure 3.34: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Red Lake. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Red Lake FWI series.

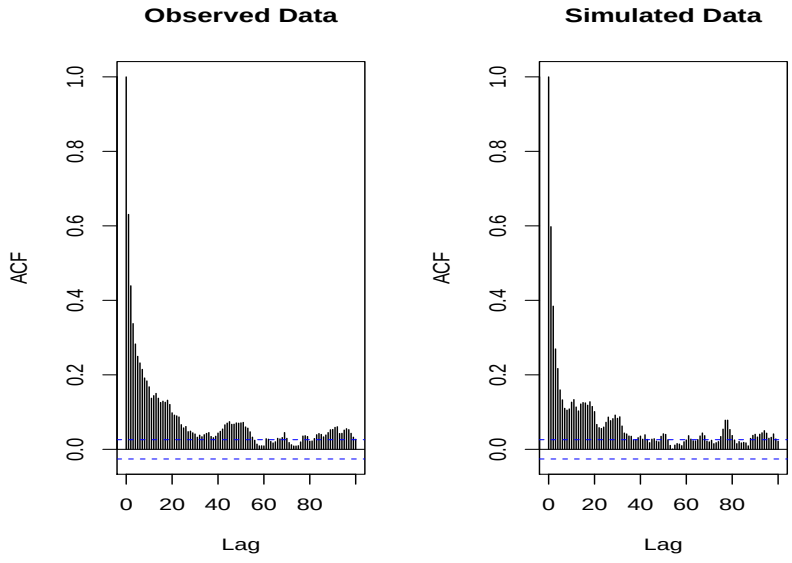

Figure 3.35: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Sioux Lookout. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Sioux Lookout FWI series.

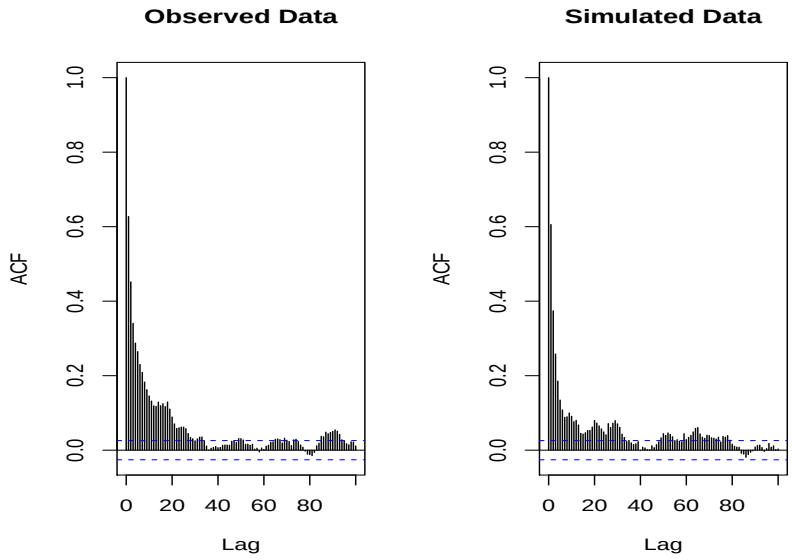

Figure 3.36: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Kenora. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Kenora FWI series.

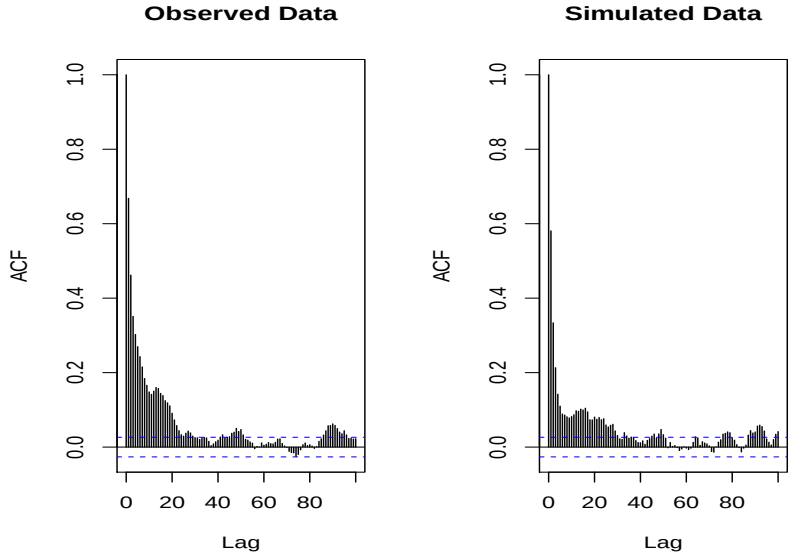

Figure 3.37: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Dryden. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Dryden FWI series.

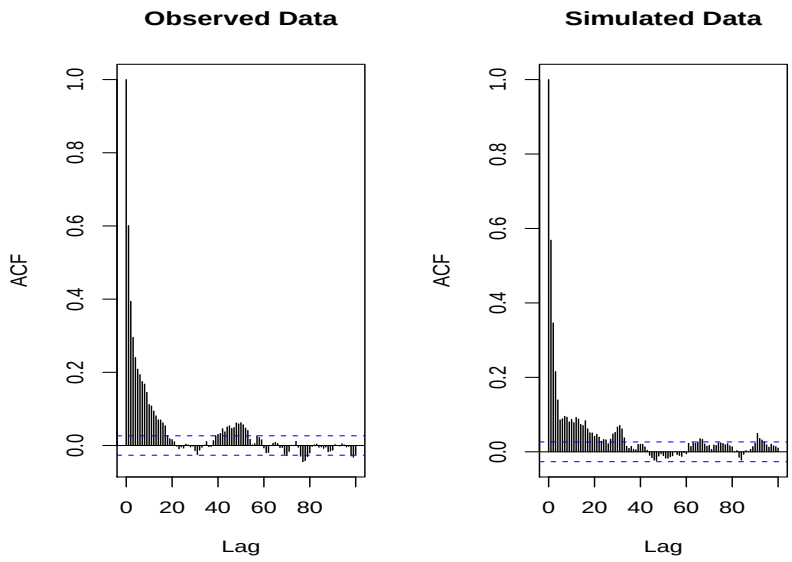

Figure 3.38: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Timmins. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Timmins FWI series.

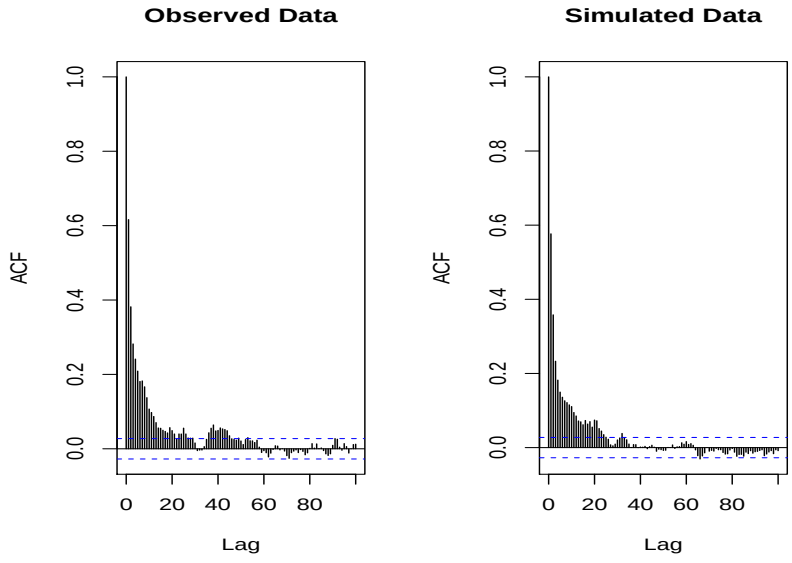

Figure 3.39: Left Panel: Sample autocorrelations for the first 100 lags for all years of FWI data from Temagami. Right Panel: Sample autocorrelations for the first 100 lags of data simulated from a sequence of seasonal models fitted to the Temagami FWI series.

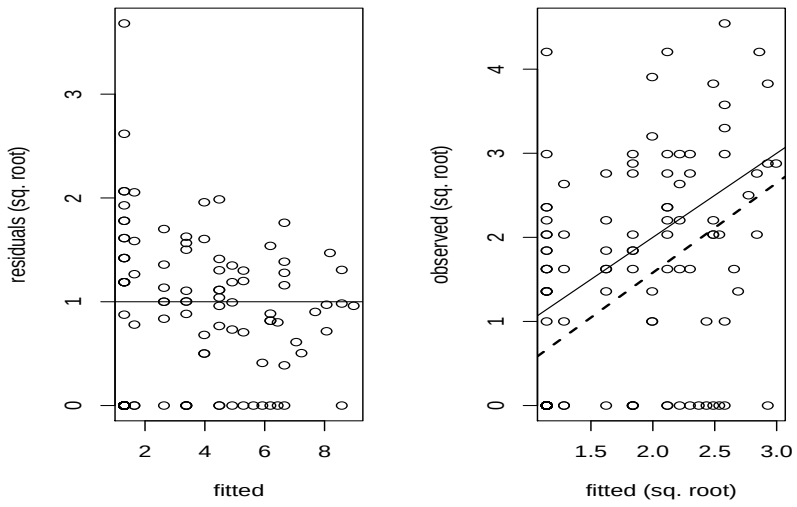

Figure 3.40: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Red Lake FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a robust smooth of this scatterplot.

where the conditional expectation is numerically evaluated from the expression given at (3.32). The lognormal probability density with parameters $\mu = 0$ and σ (estimated by maximum likelihood) is employed.

For the Weibull case, we use

$$Y_n^{1/P} \div \hat{E}[Y_n^{1/P}|Y_{n-1}]$$

where $Y_n^{1/P}$ is approximating X_n. This is much simpler to compute than $Y_n \div \hat{E}[Y_n|Y_{n-1}]$ and conveys almost the same information.

These residuals are based on ratios instead of differences, since ratios are more natural quantities to consider when working with nonnegative data. If differences were used, variation would be expected to increase with the magnitude of the conditional expectation. The use of ratios reduces, but does not completely eliminate, these kinds of spurious effects. R_n has the unfortunate side effect that when X_{n-1} is close to 0, the conditional expectation will be close to 0, but there is a relatively large probability that X_n could be nonneglible. Thus R_n would be very large in this situation. For

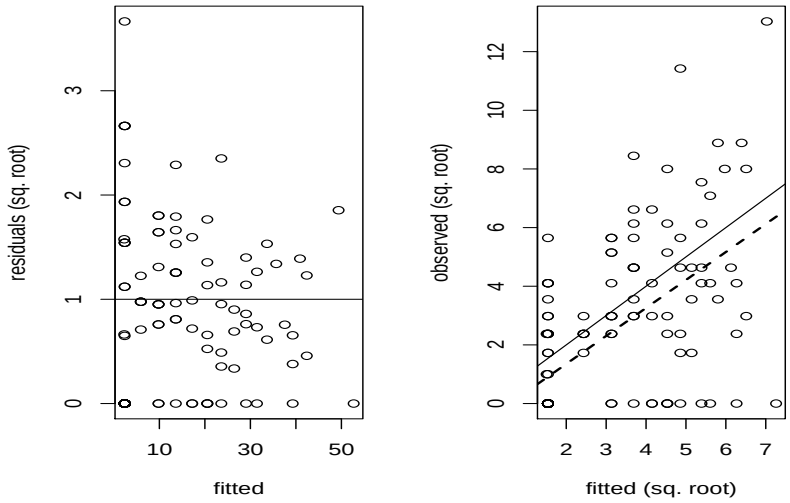

Figure 3.41: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Sioux Lookout FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a least-squares fitted line (dashed).

larger values of X_{n-1}, we expect R_n to scatter randomly about the horizontal line $y = 1$, but according to a distribution which is positively skewed.

The left panels of Figures 3.40 through 3.45 display scatterplots of the one-step-ahead predictive residuals against the corresponding conditional expectations for the 1963 FWI data for each of the six weather stations. A square root transformation has been applied to the residuals in order to reduce the distorting effects of the large residuals found at the left edge of each of the plots as well as the effects of positive skewness alluded to in the previous paragraph.

If we focus attention on the parts of the plots for which the fitted values exceed 3, we see a fairly uniform random scatter of points about the horizontal reference line. The behaviour to the left is essentially as expected. Thus, this type of residual plot is suggesting that the one-step-ahead predictions made by the model (for the 1963 data) are reasonable.

The right panels of Figures 3.40 through 3.45 display scatterplots of the FWI values (on the square root scale) against the square roots of the conditional expectations, given the previous day's FWI value, assuming the random coefficient Weibull tailed mixture model. If the model is fitting adequately, the plotted points should

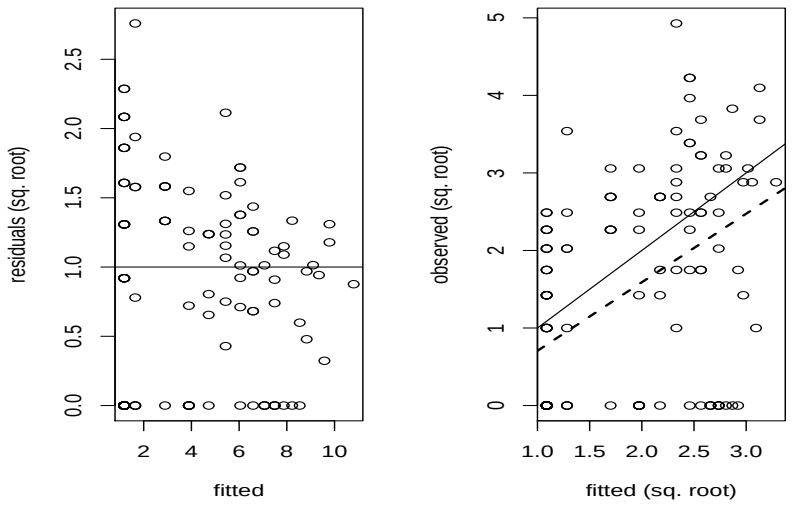

Figure 3.42: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Kenora FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a least-squares fitted line (dashed).

scatter along a 45° reference line. To assist with this check, we overlaid the least-squares fitted line relating the observed values to the fitted values. In most of the figures, the fitted line is located near the reference line, and the plotted points scatter randomly about the reference line. Therefore, by this criterion, the fitted model appears to be reasonable for the 1963 FWI data.

To check whether the probability distribution of the one-step-ahead predictions is reasonable, we can use the conditional survival functions such as

$$U(y) = P(Y_n \geq y|Y_{n-1})$$

given for the exponential case by expressions such as (3.36) and in the Weibull case by (3.38). Note that we are considering only transitions from nonzero states to nonzero states. (The next subsection considers transitions involving the nil state.)

If $\{Y'_n\}$ has the same probability law as $\{Y_n\}$, then $U(Y'_n)$ should follow a uniform distribution on the interval $[0, 1]$.

Figure 3.46 contains uniform QQ-plots of $U(\text{FWI}_n)$ for each of six weather

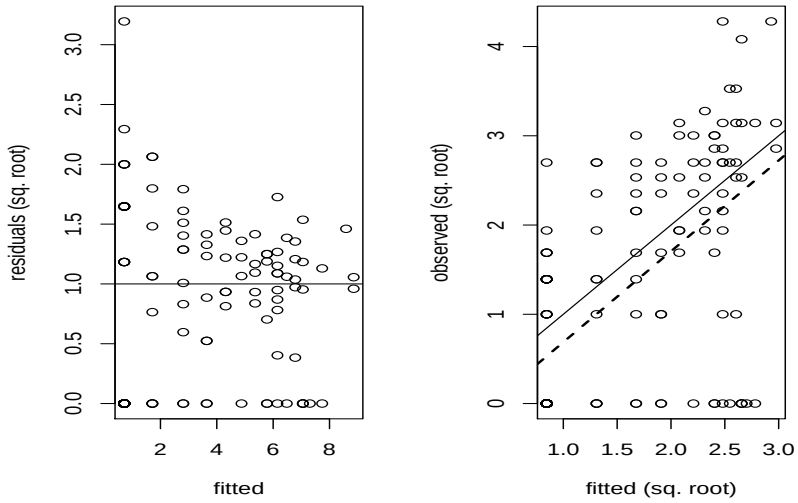

Figure 3.43: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Dryden FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a least-squares fitted line (dashed).

stations under study for 1963. The parameters used to calculate the conditional survivor functions at (3.38) were estimated from each of the 1963 FWI time series. On all six of the plots, the points are located near the 45° reference line indicating that the one-step-ahead prediction distribution is accurate.

Figure 3.47 displays the QQ-plots for all years of data at the six weather stations. In each case, the points plot very close to the reference line indicating that the predictive distribution is accurate for all years of data. In other words, one step ahead forecast distributions of FWI using the fitted model are accurate at all sites.

3.6.3 Checks on Transitions from the Nil State

When $Y_{n-1} = 0$, the tailed minification model predicts the behaviour of Y_n through the tailed variable δ_n^{P}. Under the Weibull assumption, the continuous component of the distribution of δ_n should be exponential with rate γ.

We can also check this assumption using QQ-plots restricted to those values of Y_n for which $Y_{n-1} = 0$. In those cases, $Y_n^{1/\mathrm{P}}\gamma$ should be a unit exponential random variable.

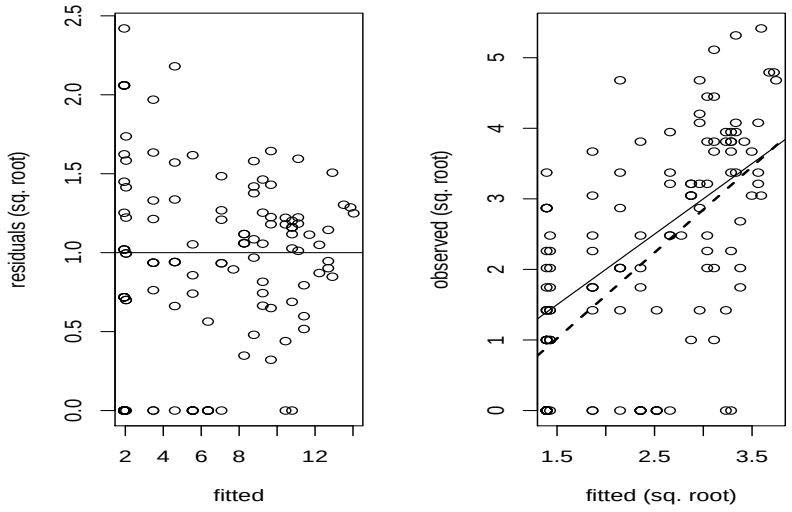

Figure 3.44: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Timmins FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a least-squares fitted line (dashed).

As an illustration, Figure 3.54 displays the QQ-plots of $\delta_n = Y_n^{1/\mathrm{P}}\gamma$ using estimates of γ for each of the 3 subseasons of the 1963 FWI datasets. The points on the plots for Kenora and Temagami deviate from the 45° reference line suggesting a discrepancy between the model predictions and the data. The other 4 plots appear satisfactory, but a close look at all of the data is warranted.

In Figure 3.55, Weibull QQ-plots are plotted for all years of data at each of the weather stations, again using the seasonal estimates of γ and P for each year. This set of plots suggests more serious trouble than was indicated on the 1963 plots. The plotted points depart from the reference line, substantially in some cases.

Thus, these plots indicate that the proposed model fails to capture this facet of the observed data. It is important to note that the pattern of departure is similar from station to station. This suggests a possible remedy: to allow the Weibull parameter P associated with δ_n to differ from the Weibull parameter associated with ε_n; the same adjustment can be used at all stations.

Figure 3.56 displays the QQ-plots of $Y_n^{1/\mathrm{P}}\gamma$ using the same estimates of γ used in Figure 3.55, but with P replaced by $\mathrm{P} \times 0.9$. The resulting plotted points lie much closer to the reference line indicating a much improved model fit. Thus, we are led

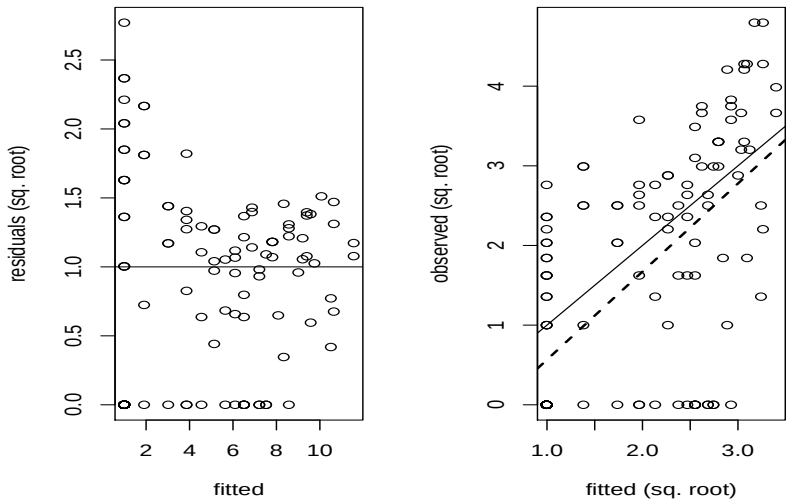

Figure 3.45: Left Panel: One-step-ahead predictive residuals (ratios $X_n/E[X_n|X_{n-1}]$) plotted against fitted values (conditional expectations of X_n, given X_{n-1}, for Temagami FWI data for 1963. Right Panel: Observed FWI values (on square root scale) plotted against the fitted values. A 45° reference line has been added, together with a least-squares fitted line (dashed).

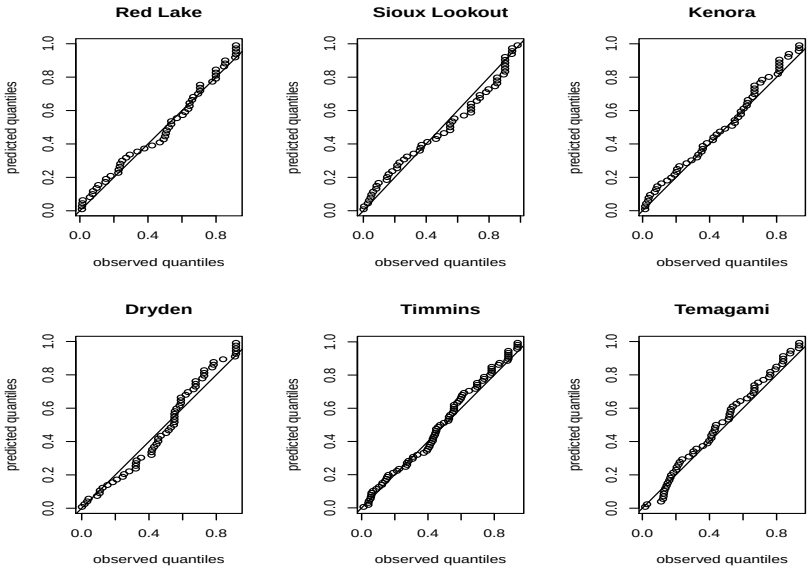

Figure 3.46: One step prediction residual QQ-plots for 1963 FWI data at all 6 stations, based on a model which ignores seasonality.

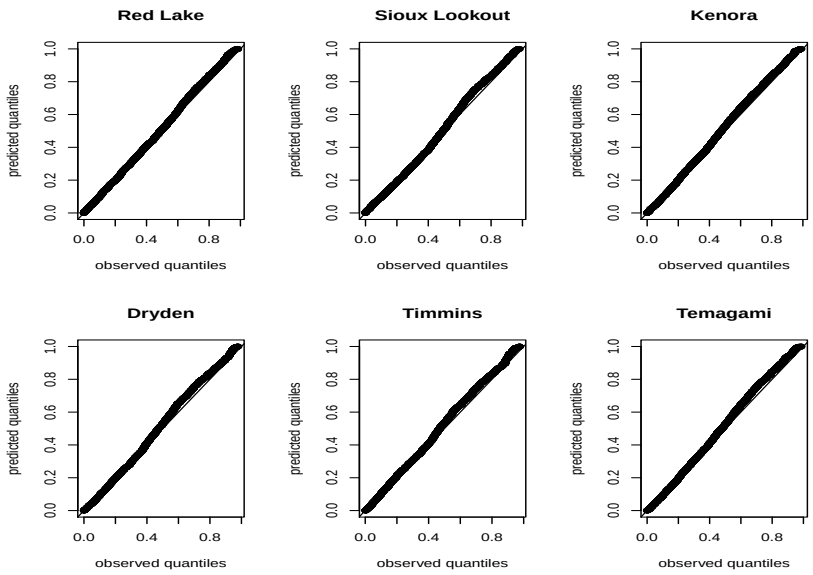

Figure 3.47: One step prediction residual QQ-plots for all years of data at all 6 stations, based on a model which ignores seasonality.

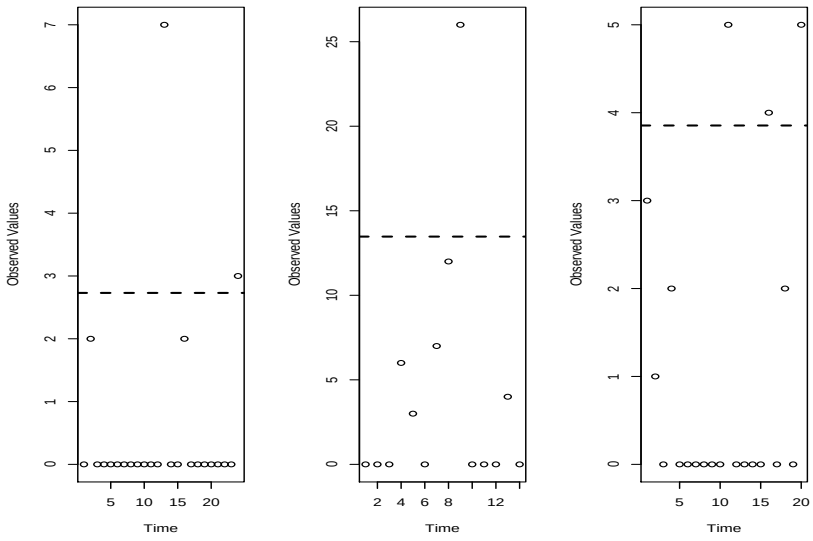

Figure 3.48: Transitions from the nil state for 1963 FWI data from Red Lake plotted against observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

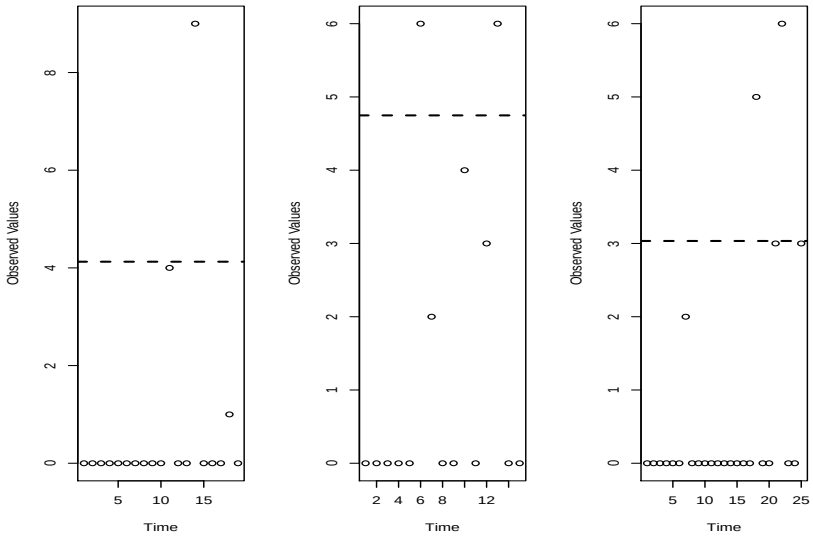

Figure 3.49: Transitions from the nil state for 1963 FWI data from Sioux Lookout plotted against observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

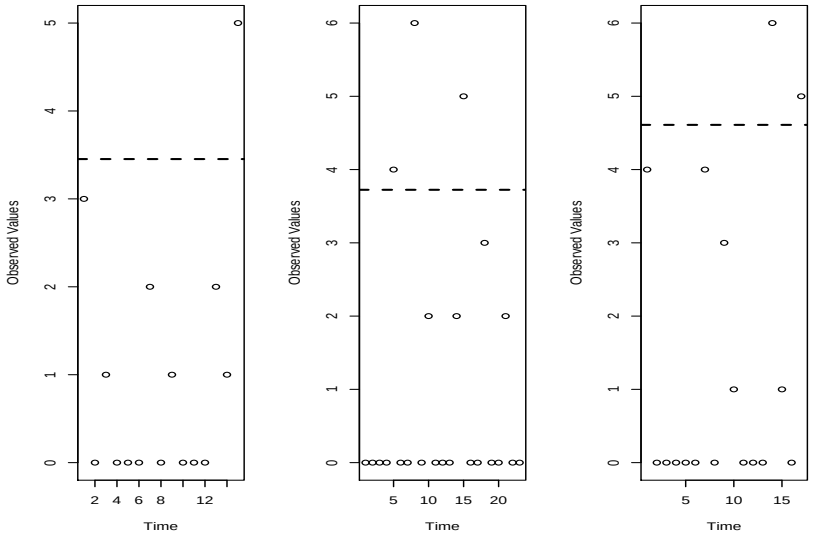

Figure 3.50: Transitions from the nil state for 1963 FWI data from Kenora plotted against observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

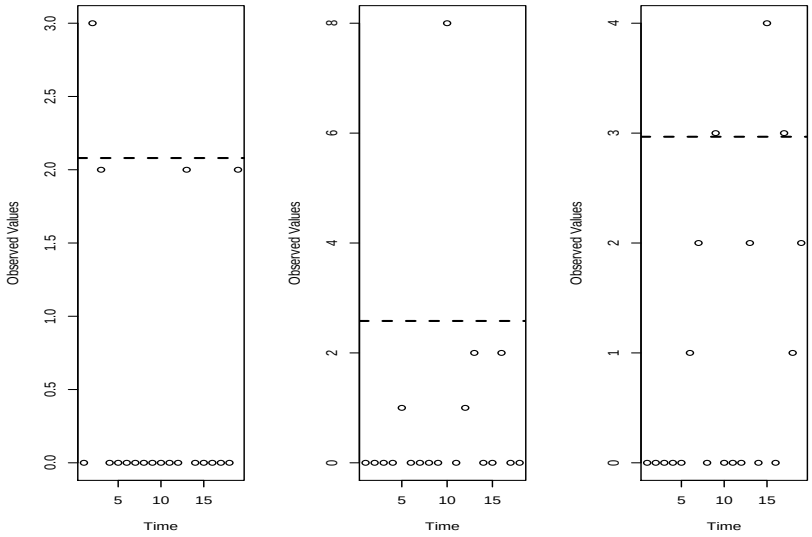

Figure 3.51: Transitions from the nil state for 1963 FWI data from Dryden plotted against observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

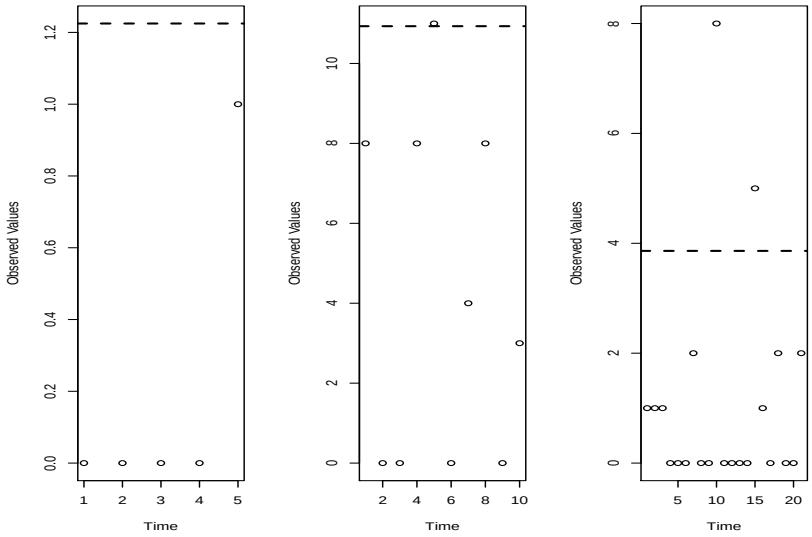

Figure 3.52: Transitions from the nil state for 1963 FWI data from Timmins plotted against observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

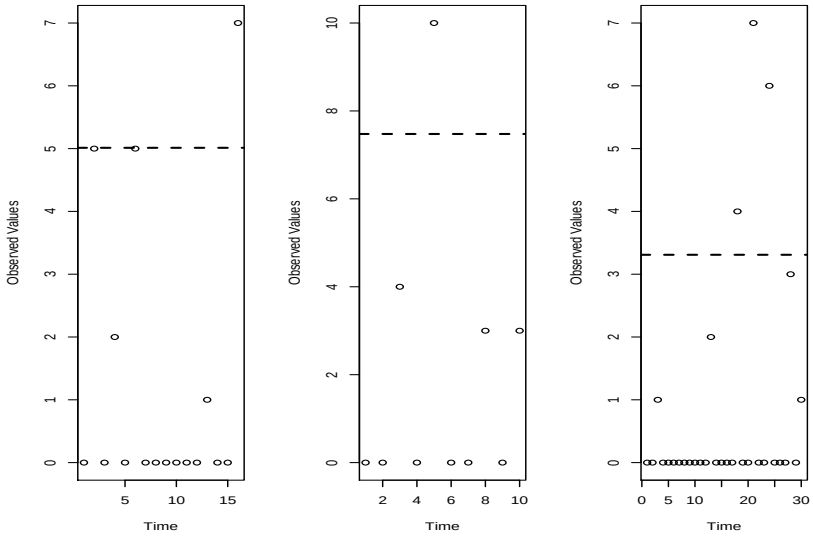

Figure 3.53: Transitions from the nil state for 1963 FWI data from Temagami observation number, with model-based 90th percentile overlaid as a dashed horizontal line. Left panel: subseason 1; middle panel: subseason2; right panel: subseason 3.

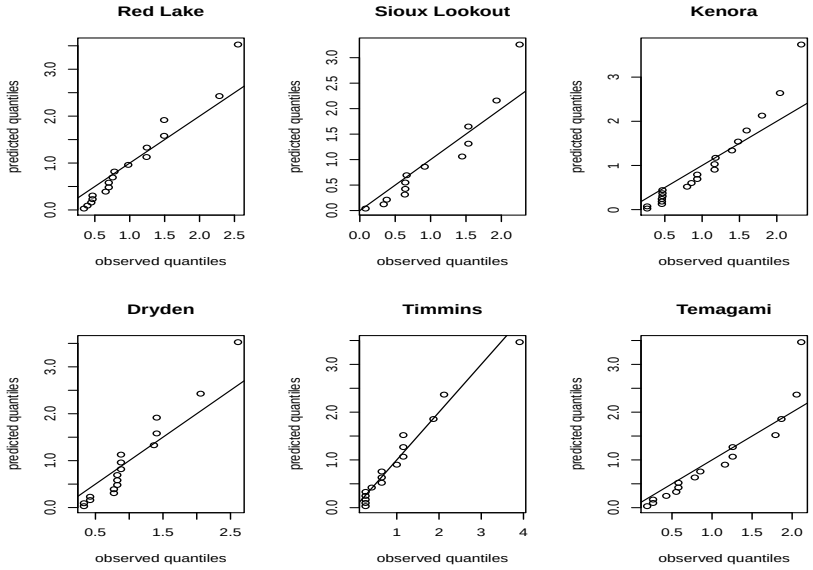

Figure 3.54: Weibull tailed QQ-plots for transitions from the nil state for 1963 FWI data from all 6 weather stations under study.

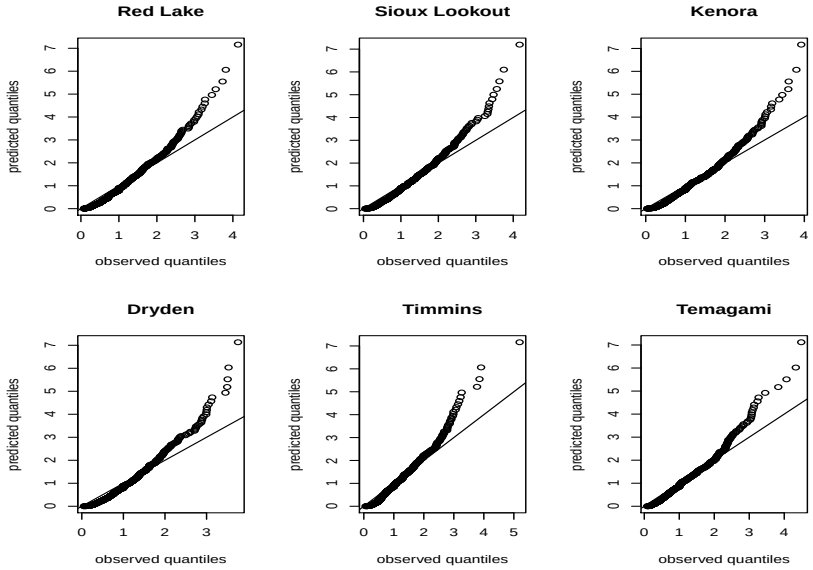

Figure 3.55: Weibull tailed δ_n QQ-plots for transitions from the nil state for all years of FWI data from all 6 weather stations under study.

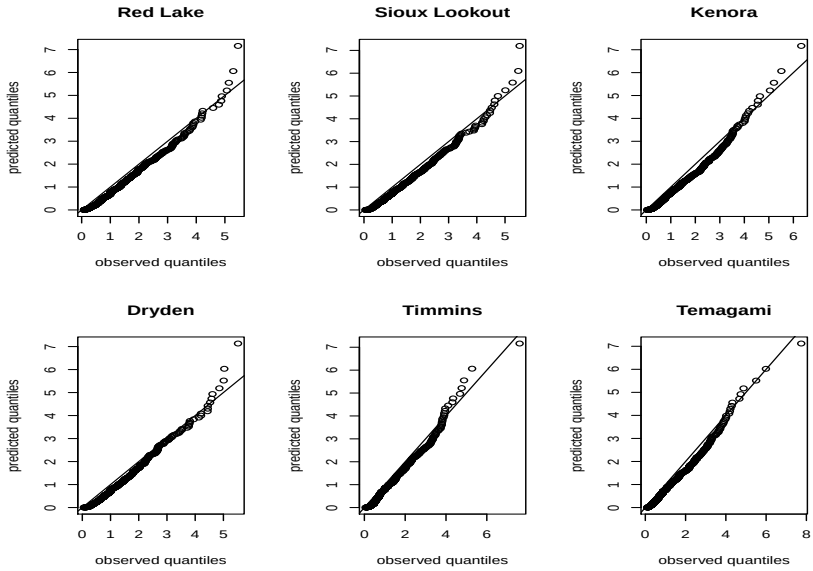

Figure 3.56: Weibull tailed δ_n^D QQ-plots for transitions from the nil state for all years of FWI data from all 6 weather stations under study. The parameter $D = 0.9$.

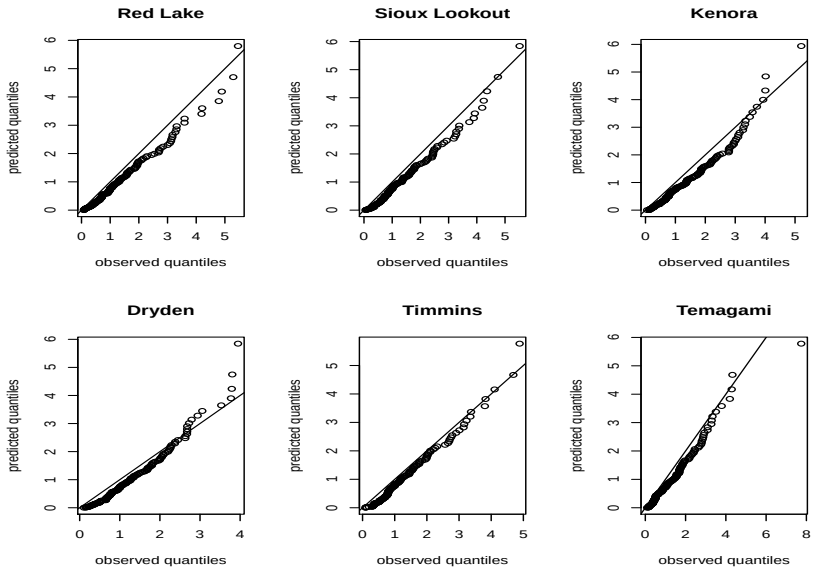

Figure 3.57: Weibull tailed δ_n^{D} QQ-plots for transitions from the nil state in subseason 1 for all years of FWI data from all 6 weather stations under study. The parameter $\mathrm{D} = 0.9$.

to the model

$$Y_n = X_n^{\mathrm{P}}$$

where

$$X_n = \begin{cases} (\alpha_n + 1)\min(X_{n-1}, \alpha_n^{-1}\varepsilon_n), & \text{if } X_{n-1} > 0. \\ \delta_n^{\mathrm{D}}, & \text{if } X_{n-1} = 0. \end{cases} \tag{3.39}$$

The new parameter D could be estimated by maximum likelihood or by inspection of QQ-plots such as displayed in Figure 3.56. These plots suggest that we could take $\widehat{\mathrm{D}} = 0.9$.

Figures 3.57 to 3.59 show the δ_n QQ-plots broken down by subseason, again for all stations under consideration. These plots indicate that there is no apparent difference in the distribution of δ_n according to season, since in all cases, the plotted points line up near the reference line in a satisfactory manner. This confirms that the same adjustment $\mathrm{D} = 0.9$ can be used for all subseasons.

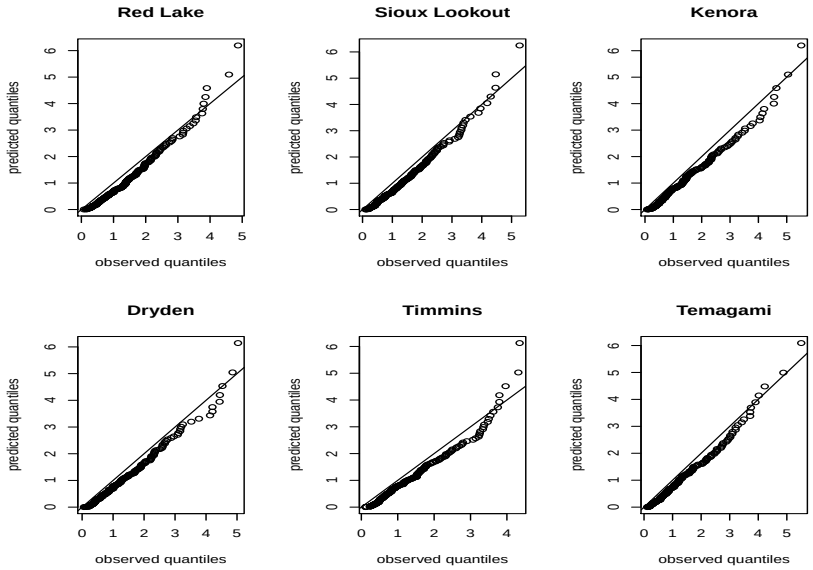

Figure 3.58: Weibull tailed QQ-plots for transitions from the nil state in subseason 2 for all years of FWI data from all 6 weather stations under study. The parameter D = 0.9.

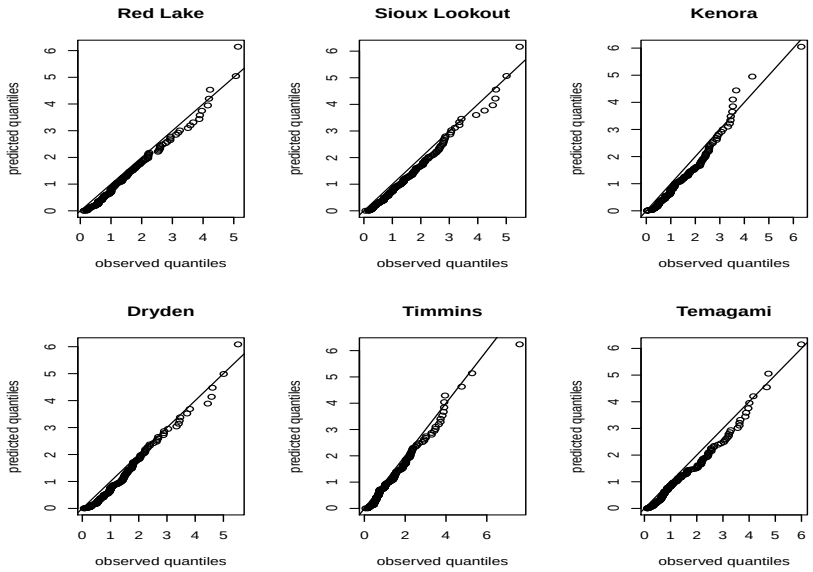

Figure 3.59: Weibull tailed δ_n^{D} QQ-plots for transitions from the nil state in subseason 3 for all years of FWI data from all 6 weather stations under study. The parameter D = 0.9.

3.7 Applying the Random Coefficient Minification Model

3.7.1 Calculation of Survival Probabilities

The model that we have developed for the Fire Weather Index can be used in many ways. A particularly important use is in the calculation of probabilities of the occurrence of large values of the FWI, since those occurrences should be associated with increased wildfire activity.

When we considered the fixed coefficient minification model, we saw that it was possible to write an explicit expression (3.13) for the stationary probability $P(X_n \geq x)$. We also noted that a recursive solution of the integral equation (3.12) is often faster and more accurate than a direct approach.

In the random coefficient case, a direct approach to finding $P(X_n \geq x)$ is possible, in principle, but it could involve a large number of iterated integrals. It appears that this approach would be computationally expensive. Instead, we focus on a numerical approximation to the equation (3.28) for $P(X_n \geq x)$, but because of the integration, a recursive method is not available.

In order to calculate the numerical approximation to $P(X_n \geq x)$, it will be useful to make a change of variable first. For any $x > 0$, let $S(x) = P(X_n \geq x)$, and let $\beta = x(\alpha + 1)^{-1}$ in (3.28), then

$$S(x) = (1 - p_\varepsilon) \int_0^x x\beta^{-2} f(x\beta^{-1} - 1) S(\beta) e^{-\lambda\beta(x\beta^{-1}-1)}\, d\beta + p_{\mathrm{x}}(1 - p_\delta)e^{-\gamma x}.$$

Recall that $f(\alpha)$ is the probability density function of the random coefficient.

When $x = 0$, the above change-of-variable transformation is not valid. Instead, we note that, from (3.28), the following is true:

$$S(0) = (1 - p_\delta)p_{\mathrm{x}} + (1 - p_\varepsilon)(1 - p_{\mathrm{x}}) = \frac{1 - p_\delta}{1 - p_\delta + p_\varepsilon} = 1 - \frac{p_\varepsilon}{1 - p_\delta + p_\varepsilon} = 1 - p_{\mathrm{x}}.$$

The following algorithm can then be used in the numerical calculation of $S(x_0)$:

1. Set $\Delta = \frac{x_0}{N}$ for some large N.
2. Set $C_1 = S(0)$.

3. Set $\widehat{S(x)} = C_1 I_{\{[0,\Delta]\}}(x)$. (Note that this is a functional assignment.)

4. For $j = 1, 2, \ldots, N$, set

   (a)

$$C_{j+1} = (1-p_\varepsilon) \int_0^{j\Delta} j\Delta\beta^{-2} f(j\Delta\beta^{-1}-1)\widehat{S(\beta)} e^{-\lambda j\Delta+\lambda\beta}\, d\beta + p_{\mathrm{x}}(1-p\delta)e^{-\gamma\Delta j}$$

   (b)

$$\widehat{S(x)} = \widehat{S(x)} + C_{j+1} I_{\{[j\Delta,(j+1)\Delta]\}}(x)$$

5. Return $C_{N+1} = \widehat{S(N\Delta)} = \widehat{S(x_0)}$.

The D adjustment can be handled by replacing step 4.(a) by

(a')

$$C_{j+1} = (1-p_\varepsilon) \int_0^{j\Delta} j\Delta\beta^{-2} f(j\Delta\beta^{-1}-1)\widehat{S(\beta)} e^{-\lambda j\Delta+\lambda\beta}\, d\beta + p_{\mathrm{x}}(1-p\delta)e^{-\gamma(\Delta j)^{1/\mathrm{D}}}$$

The corresponding Weibull-type probabilities are easily obtained through

$$P(Y_n \geq y) = P(X_n \geq y^{1/\mathrm{P}}),$$

where P is the appropriate power.

3.7.2 Calculation of Extreme Probabilities

Because of the structure of the minification model, it is particularly easy to calculate the probability of variables such as

$$Z_n = \min(X_n, X_{n-1}, \ldots, X_{n-J}).$$

To calculate $P(Z_n \geq z)$, proceed as follows:

$$P(Z_n \geq z) = P(X_{n-J} \geq z, X_{n-J+1} \geq z, \ldots, X_n \geq z)$$

$$= P(X_{n-J} \geq z, \ldots, X_{n-1} \geq z, X_{n-1} \geq z(\alpha_n + 1)^{-1}, \varepsilon_n \geq \alpha_n z(\alpha_n + 1)^{-1})$$

Note that $z > z(\alpha_n + 1)^{-1}$ is redundant. so that

$$P(Z_n \geq z) = P(Z_n \geq z)P(\varepsilon_n \geq \alpha_n z(\alpha_n + 1)^{-1})$$

Proceeding inductively, we have

$$\begin{aligned} P(Z_n \geq z) &= P(X_{n-J} \geq z)P(\varepsilon_{n-J+1} \geq \alpha_{n-J+1} z(\alpha_{n-J+1} + 1)^{-1}) \ldots \\ &\quad \ldots P(\varepsilon_n \geq \alpha_n z(\alpha_n + 1)^{-1}) \\ &= P(X_n \geq z)(P(\varepsilon \geq \alpha z(\alpha + 1)^{-1}))^{J-1} \end{aligned} \tag{3.40}$$

using the stationarity properties of $\{X_n\}$, $\{\varepsilon_n\}$ and $\{\alpha_n\}$.

We have seen in the previous subsection how to numerically approximate $P(X_n \geq z)$. Calculation of the other factor in (3.40) depends on evaluation of the integral

$$P(\varepsilon \geq \alpha z(\alpha + 1)^{-1}) = \int_0^\infty (1 - p_\varepsilon) e^{-\frac{\alpha\lambda z}{\alpha+1}} f(\alpha)\, d\alpha$$

Then (3.40) becomes

$$P(Z_n \geq z) = P(X_n \geq z) \left(\int_0^\infty (1 - p_\varepsilon) e^{-\frac{\alpha\lambda z}{\alpha+1}} f(\alpha)\, d\alpha \right)^{J-1}$$

3.7.3 Calculation of Model-Based FWI Exceedance Levels

The probability expressions developed in the previous two subsections have potential use in the wildfire management context. With such formulas, it is possible to deter-

mine location-specific percentiles for the FWI which could be used to set cut-offs for local fire danger.

Fire activity may not increase because of a single day of elevated FWI. Instead, it may take two or more days of increased FWI to generate fires. Thus, the probabilities of minima can also be used to set local fire danger cut-offs.

We briefly demonstrate how the percentile cut-offs might be determined at the six weather station locations. To do this, we need parameter values that are "typical" for a station. A simple approach is to compute the medians of all annual parameter estimates at each weather station. Table 3.9 contains the medians of the random coefficient minification process parameter estimates.

	P	p_ε	p_δ	λ	γ	σ
Red Lake	1.00	0.18	0.62	0.12	0.31	1.17
Sioux Lookout	0.99	0.18	0.64	0.11	0.34	1.20
Kenora	0.95	0.17	0.59	0.11	0.27	1.21
Dryden	0.97	0.18	0.65	0.11	0.30	1.26
Timmins	1.03	0.17	0.60	0.15	0.39	1.18
Temagami	1.01	0.20	0.62	0.12	0.32	1.24

Table 3.9: Medians of parameter estimates for the random coefficient Weibull tailed mixture model fit to each of the six Ontario weather stations under study.

Tables 3.10 through 3.15 display the probabilities of exceeding the given values of the FWI listed in the left-most column, for each of the weather stations. The second column lists the probability of exceeding the FWI value for one day, the third column lists the probability of exceeding the FWI value for two consecutive days, and so on. The probability calculations are based on the parameter estimates listed in Table 3.9.

To illustrate the use of the tables, we can find the probability that the minimum FWI value on 4 consecutive days exceeds 15.0 under typical conditions at Red Lake; according to Table 3.10, it is 0.00405, a relatively rare event.

Incorporating the D adjustment into the δ_n component of the minification model can be shown to have little effect on these probability calculations. Table 3.16 is a replicate of Table 3.15 where the only difference in its calculation is to use the D = 0.9 adjustment. It can be seen that the entries in the two tables are very similar. Thus, for the purpose of calculating these kinds of exceedance probabilities, the simpler model (without adjustment) is sufficient for practical purposes.

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.571	0.441	0.3401	0.26248	0.20259
3.000	0.414	0.286	0.1969	0.13575	0.09359
5.000	0.306	0.189	0.1166	0.07202	0.04447
7.000	0.229	0.127	0.0705	0.03910	0.02168
9.000	0.174	0.087	0.0434	0.02165	0.01081
11.000	0.133	0.060	0.0271	0.01220	0.00550
13.000	0.103	0.042	0.0171	0.00699	0.00285
15.000	0.080	0.030	0.0109	0.00405	0.00150
17.000	0.062	0.021	0.0071	0.00238	0.00080
19.000	0.048	0.015	0.0046	0.00142	0.00044
21.000	0.038	0.011	0.0030	0.00085	0.00024
23.000	0.030	0.008	0.0020	0.00052	0.00013
25.000	0.023	0.006	0.0013	0.00032	0.00007
27.000	0.019	0.004	0.0009	0.00019	0.00004
29.000	0.015	0.003	0.0006	0.00012	0.00002

Table 3.10: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Red Lake district.

3.7.4 Fire Occurrence Prediction

The minification process model may provide new ways of quantifying the use of FWI in the prediction of fire flaps, events usually associated with multiple fires.

In this subsection, we demonstrate through a few particular examples how FWI exceedance is associated with fire activity. Simple graphics can be used to visualize this association. The minification model gives us a way of calculating the probability of such events.

In our first example (Figure 3.60), a dot plot is used to identify dates on which the FWI exceeded 21 at Red Lake, for each of the years from 1976 through 1981. The model-based probability of such occurrences is 0.038, according to Table 3.10.

Red circles are used to indicate days in which fires were reported in the Red Lake district. A small amount of vertical jittering was used in order to visualize multiple fires on the same date.

What is noteworthy in this figure is the large concentrations of red circles at several locations indicating the level of prevalence of fire activity in the district. There are several concentrations of fire events that are not near black dots, indicating that the FWI values were not above 21 very often even when fires eventually occurred. Ideally, one would like to see at least one black dot immediately preceding a concentration of red circles on this kind of plot. This would be an indication that FWI

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.556	0.430	0.3323	0.25694	0.19868
3.000	0.400	0.277	0.1920	0.13307	0.09224
5.000	0.294	0.183	0.1140	0.07095	0.04417
7.000	0.220	0.124	0.0692	0.03881	0.02175
9.000	0.168	0.085	0.0429	0.02169	0.01097
11.000	0.129	0.059	0.0270	0.01235	0.00565
13.000	0.100	0.041	0.0172	0.00714	0.00297
15.000	0.078	0.029	0.0111	0.00419	0.00158
17.000	0.061	0.021	0.0072	0.00248	0.00085
19.000	0.048	0.015	0.0047	0.00149	0.00047
21.000	0.038	0.011	0.0031	0.00090	0.00026
23.000	0.030	0.008	0.0021	0.00055	0.00015
25.000	0.023	0.006	0.0014	0.00034	0.00008
27.000	0.018	0.004	0.0009	0.00021	0.00005
29.000	0.015	0.003	0.0006	0.00013	0.00003

Table 3.11: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Sioux Lookout district.

levels above 21 are predictive of the occurrence of fire flaps. At the same time, too many black dots appearing without being followed by a red circle would indicate a false alarm.

Our conclusion from this analysis is that predicting fire flaps from the observation that the FWI exceeds 21 does not work well at Red Lake.

In Figure 3.61, the same fire ignitions are plotted as in the previous figure, but the black dots are located at the dates on which 3 consecutive days with an FWI value above 13 have occurred. Under typical conditions, the probability of such an event is .0171. According to the plots, these typically rare events are occurring with relatively high frequency and appear to be associated more strongly with fire activity. This analysis is indicating that it may be better to use a warning based on 3 days worth of FWI measurements than only a single day.

Continuing with the same set of fire ignitions, Figure 3.62 displays the dates when the minimum FWI over the previous 5 days exceeds 11. Under typical conditions at Red Lake, Table 3.10 suggests that the probability of this event is 0.00550. In most cases, the occurrence of such events seems to be associated with fire activity, with the exception of the plot for 1981 which shows several dates in which this type of extreme FWI event is not associated with fires.

For a follow-up example, we repeated the above graphical analysis on the Temagami district for the same time period. Note that Temagami is in a very differ-

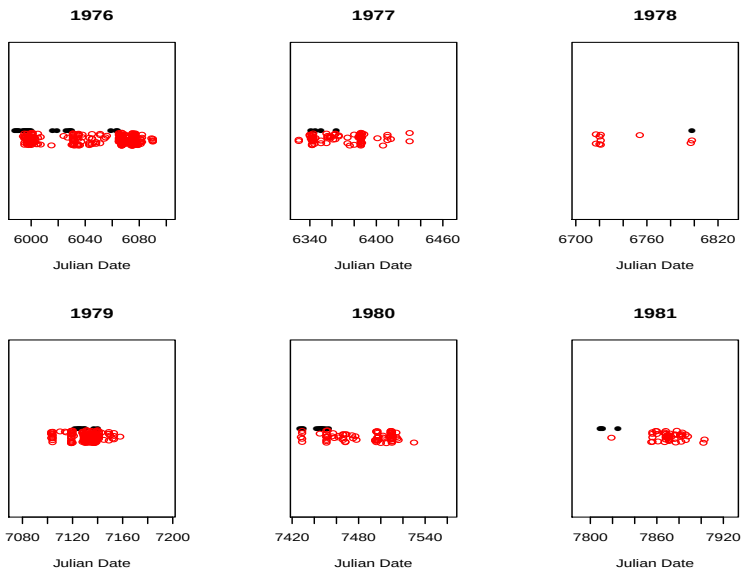

Figure 3.60: FWI exceedances and fire occurrence in the Red Lake district, for 1976 through 1981. Black dots: days when the FWI exceeds 21; red circles: days when fires have started.

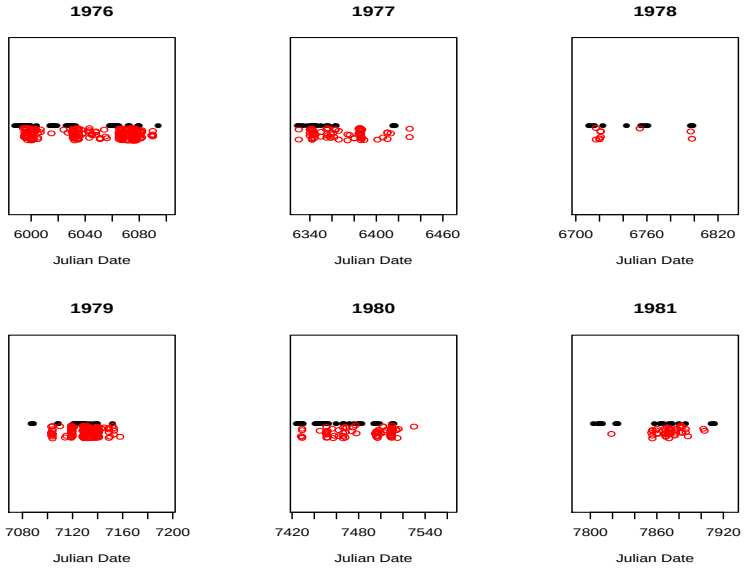

Figure 3.61: FWI exceedances and fire occurrence in the Red Lake district, for 1976 through 1981. Black dots: days when the FWI has exceeded 13 for at least 3 days; red circles: days when fires have started.

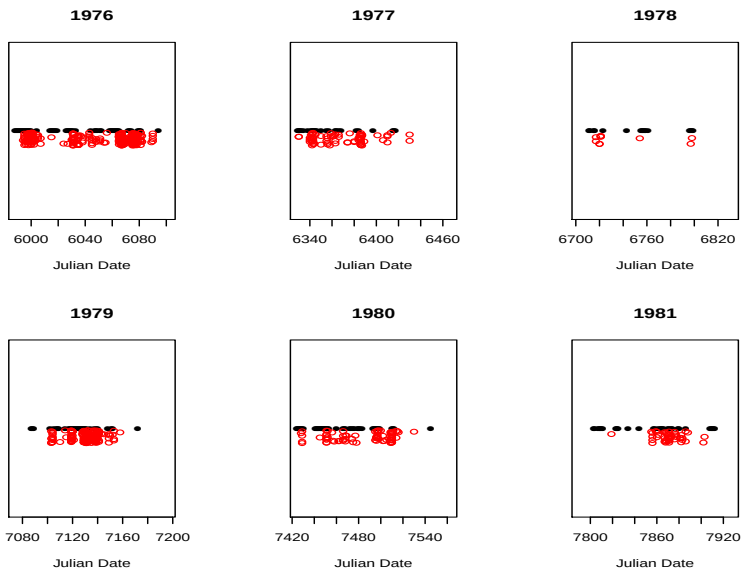

Figure 3.62: FWI exceedances and fire occurrence in the Red Lake district, for 1976 through 1981. Black dots: days when the FWI has exceeded 11 for at least 5 days; red circles: days when fires have started.

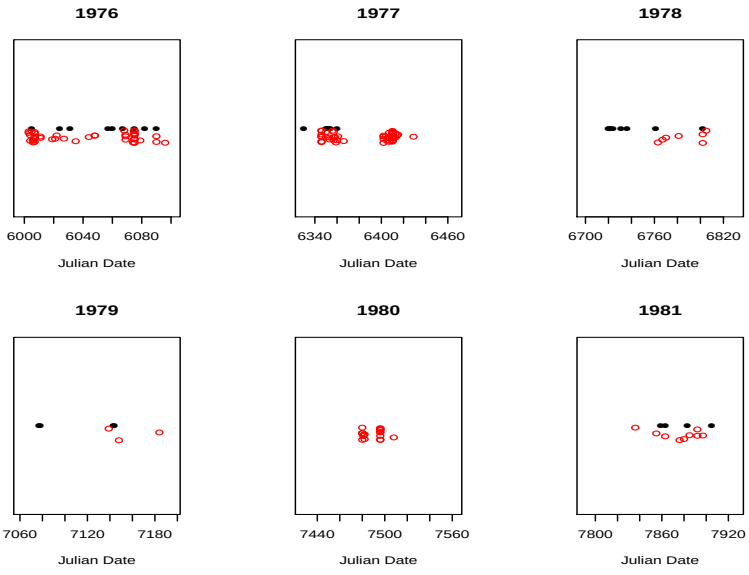

Figure 3.63: FWI exceedances and fire occurrence in the Temagami district, for 1976 through 1981. Black dots: days when the FWI exceeds 21; red circles: days when fires have started.

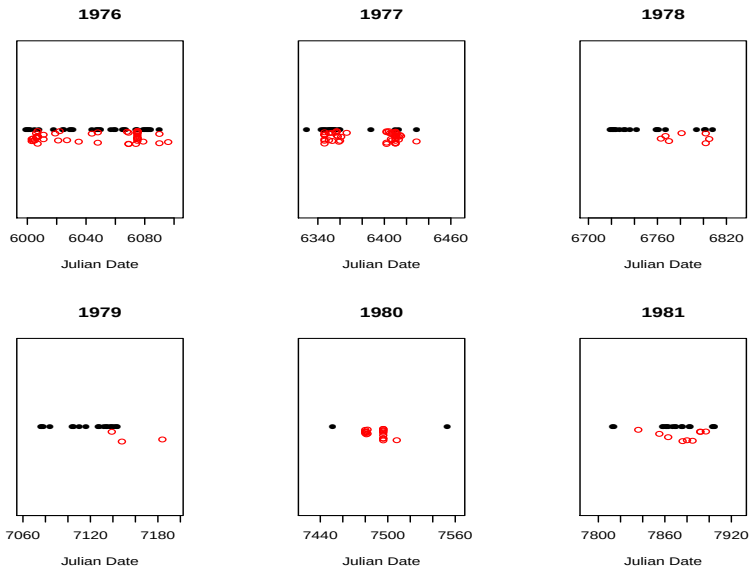

Figure 3.64: FWI exceedances and fire occurrence in the Temagami district, for 1976 through 1981. Black dots: days when the FWI has exceeded 13 for at least 3 days; red circles: days when fires have started.

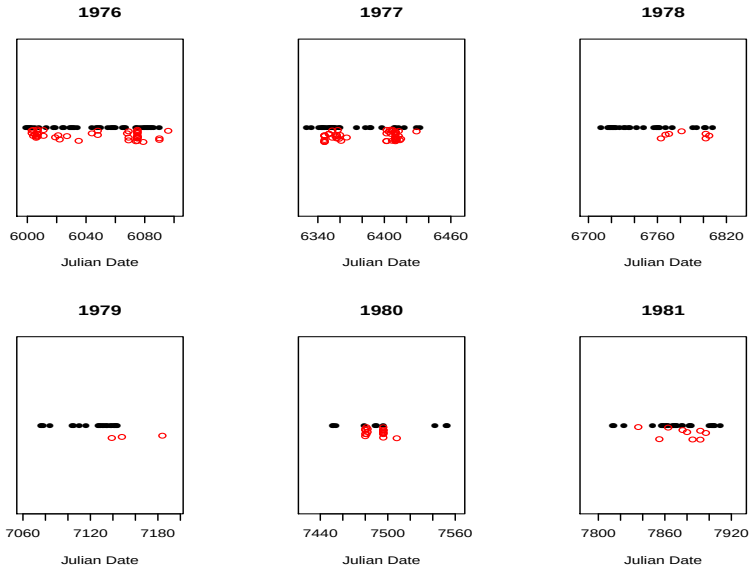

Figure 3.65: FWI exceedances and fire occurrence in the Temagami district, for 1976 through 1981. Black dots: days when the FWI has exceeded 11 for at least 5 days; red circles: days when fires have started.

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.605	0.475	0.3726	0.29227	0.22930
3.000	0.447	0.314	0.2202	0.15465	0.10860
5.000	0.331	0.208	0.1306	0.08205	0.05153
7.000	0.247	0.139	0.0782	0.04399	0.02474
9.000	0.186	0.094	0.0474	0.02388	0.01204
11.000	0.141	0.064	0.0290	0.01313	0.00595
13.000	0.107	0.044	0.0179	0.00730	0.00298
15.000	0.082	0.030	0.0111	0.00411	0.00152
17.000	0.063	0.021	0.0070	0.00234	0.00078
19.000	0.048	0.015	0.0044	0.00134	0.00041
21.000	0.037	0.010	0.0028	0.00078	0.00021
23.000	0.029	0.007	0.0018	0.00046	0.00011
25.000	0.022	0.005	0.0012	0.00027	0.00006
27.000	0.017	0.004	0.0008	0.00016	0.00003
29.000	0.013	0.003	0.0005	0.00010	0.00002

Table 3.12: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Kenora district.

ent part of the province from Red Lake so there is no prior reason to expect the same kind of fire and weather activity.

Looking at Figure 3.63, we see several concentrations of red circles without any black dots nearby, as well as several instances of black dots not being followed by red circles. This means that looking at days when the FWI exceeds 21 is again not a very predictor of fire activity.

The black dot plot on Figure 3.64 shows the dates when the minimum FWI on 3 consecutive days exceeds 13 (an event with probability 0.0159 under typical conditions, according to Table 3.15). Although there are several fire flaps that are preceded by one or more black dots, there are some fires, notably in the 1980 plot, which are not predicted by three consecutive days of FWI above 13. In addition, there are some false alarms in 1978 and 1979.

Finally, Figure 3.65 shows dates when the minimum FWI on 5 consecutive days exceeds 11 (an event with probability 0.00492, according to Table 3.15). In this case, the fire flap in 1980 is correctly predicted, but the numbers of false alarms in 1978 and 1979 is even larger than before.

This exploratory analysis is clearly not complete and will be the subject of further research, but it is highly suggestive of the importance of looking at the FWI as a predictor for fire activity in nontraditional ways. A formal analysis would require the study of all fires and different configurations of minimal statistics based on the

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.566	0.442	0.3445	0.26877	0.20968
3.000	0.409	0.285	0.1980	0.13774	0.09581
5.000	0.299	0.186	0.1153	0.07151	0.04437
7.000	0.222	0.123	0.0681	0.03778	0.02095
9.000	0.166	0.082	0.0409	0.02031	0.01009
11.000	0.125	0.056	0.0249	0.01109	0.00495
13.000	0.095	0.038	0.0153	0.00615	0.00247
15.000	0.072	0.026	0.0095	0.00345	0.00125
17.000	0.055	0.018	0.0060	0.00196	0.00065
19.000	0.042	0.013	0.0038	0.00113	0.00034
21.000	0.032	0.009	0.0024	0.00066	0.00018
23.000	0.025	0.006	0.0015	0.00039	0.00010
25.000	0.019	0.004	0.0010	0.00023	0.00005
27.000	0.015	0.003	0.0007	0.00014	0.00003
29.000	0.011	0.002	0.0004	0.00008	0.00002

Table 3.13: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Dryden district.

FWI, and is outside the scope of the current study. Our present interest has been in proposing the minification model and indicating how it might be used in practice.

3.7.5 Bootstrap Hypothesis Testing

With a model that gives a reasonable fit to the data, it is now possible to perform certain hypothesis tests of interest. In particular, parametric bootstrap tests can be conducted.

In this section, we will test the hypothesis that all lag 1 autocorrelations are the same from year to year against the alternative that at least one is different. This hypothesis was also tested in Han and Braun (2014), and we will use the same test statistic here. The only difference between the present approach and that one is that we will construct bootstrap samples by simulating from the fitted model instead of conducting a block bootstrap.

In order to simulate FWI time series that satisfy the null hypothesis, we use the median of all parameter estimates given in Table 3.9 as the parameter values to repeatedly generate bootstrap sample replicates.

As in Han and Braun (2014), the test statistic is calculated as the standard deviation of the annual lag 1 autocorrelation estimates of the FWI values raised to the power 0.2.

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.565	0.434	0.3327	0.25531	0.19591
3.000	0.387	0.259	0.1734	0.11608	0.07769
5.000	0.273	0.161	0.0945	0.05552	0.03263
7.000	0.197	0.102	0.0531	0.02757	0.01431
9.000	0.145	0.067	0.0307	0.01411	0.00650
11.000	0.107	0.044	0.0180	0.00741	0.00304
13.000	0.080	0.029	0.0108	0.00397	0.00146
15.000	0.060	0.020	0.0066	0.00217	0.00072
17.000	0.046	0.014	0.0040	0.00121	0.00036
19.000	0.035	0.009	0.0025	0.00068	0.00018
21.000	0.026	0.006	0.0016	0.00039	0.00010
23.000	0.020	0.005	0.0010	0.00022	0.00005
25.000	0.015	0.003	0.0006	0.00013	0.00003
27.000	0.012	0.002	0.0004	0.00008	0.00001
29.000	0.009	0.002	0.0003	0.00005	0.00001

Table 3.14: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Timmins district.

Using 500 bootstrap replicates for each weather station, and calculating the bootstrap p-value as the proportion of bootstrap samples in which the resulting test statistic exceeds the test statistic for the original sample. The p-values for the six stations are

Red Lake	Sioux Lookout	Kenora	Dryden	Timmins	Temagami
0.024	0.004	0.006	0.1	0.248	0.066

The conclusion that we draw from this set of tests is that there is clear evidence that the lag 1 autocorrelations change from year to year at Red Lake, Sioux Lookout and Kenora. The evidence is less clear at the other three stations. This result is slightly different from what was reported in Han and Braun (2014) where a confidence interval analysis indicated that the annual lag 1 autocorrelations at Dryden are not constant. Otherwise, the conclusion is the same using both the nonparametric and parametric procedures.

3.8 Conclusions and Further Work

FWI data are nonnegative and include many zeros. The FWI increases when the relative humidity is low, as the temperature rises, and with increases in wind speed. Increases in FWI could also occur when cloud cover is thin or nonexistent. This means

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.549	0.415	0.3136	0.23688	0.17895
3.000	0.395	0.267	0.1799	0.12132	0.08183
5.000	0.291	0.176	0.1063	0.06425	0.03885
7.000	0.217	0.118	0.0643	0.03497	0.01902
9.000	0.165	0.081	0.0397	0.01949	0.00957
11.000	0.126	0.056	0.0249	0.01108	0.00492
13.000	0.097	0.039	0.0159	0.00641	0.00259
15.000	0.076	0.028	0.0102	0.00377	0.00138
17.000	0.059	0.020	0.0067	0.00224	0.00075
19.000	0.046	0.014	0.0044	0.00135	0.00042
21.000	0.037	0.010	0.0029	0.00083	0.00023
23.000	0.029	0.008	0.0020	0.00051	0.00013
25.000	0.023	0.005	0.0013	0.00032	0.00008
27.000	0.018	0.004	0.0009	0.00020	0.00004
29.000	0.014	0.003	0.0006	0.00013	0.00003

Table 3.15: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Temagami district.

FWI increases are gradual but random. Sudden rain causes the FWI to decrease, often to the nil value.

At the beginning of the fire season (Subseason 1), rain events may tend to happen more often than at other times in the fire season. Thus, in the different subseasons drying rates and their associated model coefficients are different.

The histogram of FWI excluding zeros shows an exponential-like shape. Through a sequence of model-building steps, this led to our proposal of a random coefficient mixture Weibull tailed minification model:

$$Y_n = X_n^{\mathrm{P}}$$

where

$$X_n = \begin{cases} (\alpha_n + 1)\min(X_{n-1}, \alpha_n^{-1}\varepsilon_n), & \text{if } X_{n-1} > 0. \\ \delta_n, & \text{if } X_{n-1} = 0. \end{cases}$$

Here ε_n, δ_n are independent exponential tailed random variables and α_n is lognormal. For the Ontario data, we showed that a mixture Weibull tailed distribution works reasonalbly well, but seasonality should be accounted for.

This model is a Markovian process. Using techniques from ODE theory, we showed the existence and uniqueness of the stationary distribution of the homoge-

FWI	1 Day	2 Days	3 Days	4 Days	5 Days
1.000	0.547	0.413	0.3124	0.23598	0.17827
3.000	0.401	0.270	0.1823	0.12299	0.08296
5.000	0.300	0.181	0.1097	0.06633	0.04010
7.000	0.228	0.124	0.0674	0.03663	0.01992
9.000	0.174	0.086	0.0420	0.02062	0.01012
11.000	0.134	0.060	0.0266	0.01181	0.00525
13.000	0.104	0.042	0.0170	0.00686	0.00277
15.000	0.081	0.030	0.0110	0.00404	0.00149
17.000	0.064	0.021	0.0072	0.00241	0.00081
19.000	0.050	0.015	0.0047	0.00145	0.00045
21.000	0.039	0.011	0.0031	0.00089	0.00025
23.000	0.031	0.008	0.0021	0.00055	0.00014
25.000	0.024	0.006	0.0014	0.00034	0.00008
27.000	0.019	0.004	0.0010	0.00021	0.00005
29.000	0.015	0.003	0.0007	0.00014	0.00003

Table 3.16: Probabilities of exceeding a given FWI value for 1, 2, 3, 4 or 5 consecutive days in the Temagami district.

neous Markov process. Although it gives realistic simulations and accurate probability calculations, FWI time series have been shown to not exactly satisfy the properties of a homogeneous Markovian process in a recent paper by Albert-Green et al (2013). This could be due in part to seasonality, but there could also be some truly non-Markovian character to the true process. Nevertheless, the model we propose in this paper appears to fit the data well.

The model proposed in the current paper will also be subject to certain inaccuracies. While it is possible to incorporate mechanisms into the random coefficient minification model which would render it non-Markovian, the price for doing so may be more than it is worth. The current model is computationally tractable and easily simulated, providing good approximations to marginal and joint distributions of the FWI values, but the small improvement in approximation accuracy coming from a non-Markovian model may result in a model which is difficult to simulate from and for which probability calculations are prohibitively difficult. One alternative would be to incorporate further randomization by allowing p_δ and/or p_ε to become random variables, possibly with beta distributions. This is a subject for further research.

Parameter estimation for the random coefficient minification model by maximum likelihood is feasible. Several checks on the fitted model were undertaken. Comparing long run autocorrelation plots of observed data and simulated data with and without seasonality for each location, it is evident that seasonality is quite im-

portant. The seasonal ones can match most of the first 20 lags correlation with slight underestimation but much better than the nonseasonal model. The predictive residual plots provide additional evidence of the accuracy of the proposed model. Weibull tailed δ_n QQ-plots for transition probabilities from the nil state confirm the model fitting, but additional adjustment is necessary to handle nil state transitions.

Allowing the Weibull parameter to differ for the δ_n component of the model needs further investigation. The model is still a member of the general family for which we proved existence and uniqueness of a stationary distribution:

$$Y_n = \begin{cases} (\alpha_n + 1)\min(Y_{n-1}, Z_n), & \text{if } Y_{n-1} > 0. \\ W_n, & \text{if } Y_{n-1} = 0. \end{cases}$$

The adjusted model presented here has $W_n = \delta^{\text{PD}}$. The value of D was selected informally. Likelihood estimation of the D adjustment should be considered in the future.

We considered applications of the random coefficient minification model to testing and evaluating probabilities of extreme events. Likelihood ratio testing should be considered in the future; the asymptotic distribution of the likelihood ratio will require careful study before it can be applied in practice. The association between fire events and extreme FWI behaviour warrants additional research, since we have provided evidence that consideration of runs of several days of moderate FWI values could be more closely associated with fires than single days of high FWI values. Our approach has been informal. A systematic formal approach is needed before a firm conclusion can be reached.

BIBLIOGRAPHY

[1] Adke, S. R. and Balakrishna, N. (1992). Estimation of the mean of some stationary Markov sequences. *Comm. Statist.–Theory Meth*, **21**, 137–159.

[2] Albert-Green, A., Braun, W.J., Martell, D.L., and Woolford, D.G. (2013). Visualization Tools for Assessing the Markov Property: Sojourn Times in the Ontario Fire Weather Index. *Environmetrics*, doi: 10.1002/env.2237

[3] Balakrishna, N. and Jacob, T.M. (2003). Parameter estimation in minification processes. *Commun. Statist. – Theory Meth*, **32**, 2139–2152.

[4] Chernick, M. R., Daley, D. J., Littlejohn, R. P. (1988). A time-reversibility relationship between two Markov chains with exponential stationary distributions. *J. Appl. Prob.* **25**, 418–422.

[5] Forestry Canada Fire Danger Group (1992). Development and structure of the Canadian Forest Fire behaviour Prediction System. Forestry Canada, Ottawa, Ontario. Information Report ST-X-3.

[6] Fujioka, F.M. and Tsou, T. (1985). Probability modelling of a fire weather index. In *Proceedings of the 8th Conference on Fire and Forest Meteorology.* (Edited by L.R. Donoghue and R.E. Martin) Society of American Foresters, 239–243.

[7] Gaver, D.G. and Lewis, P.A.W. (1980). First-order autoregressive Gamma sequences and point processes. *Adv. Appl. Prob.*, **12**, 727–745.

[8] Han, L. and Braun, W.J. (2014). Block bootstrap calibration with application to the fire weather index. To appear in *Comm. Statist. – Simulation and Computation.*

[9] Lawrance, A.J. and Lewis, P.A.W. (1977). An exponential moving average sequence and point process, EMA(1). *J. Appl. Prob.*, **14**, 98–113.

[10] Lawrance, A.J. and Lewis, P.A.W. (1980). The exponential autoregressive-moving average EARMA(p,q) process. *J. Roy. Statist. Soc. Ser. B*, **42**, 150–161.

[11] Lewis, P.A.W., Mckenzie, E.D. (1991). Minification processes and their transformations. *J. Appl. Prob.*, **28**, 45–57.

[12] Littlejohn, R.P. (1994). A reversibility relationship for two Markovian time series models with stationary exponential tailed distribution. *J. Appl. Prob.*, **31**, 575–581.

[13] Marshall, A.W. and Olkin, I. (1967) A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, **62**, 30–44.

[14] Martell, D.L. (1999). A Markov chain model of day to day changes in the Canadian forest fire weather index. *International Journal of Wildland Fire*, **9**, 265–273.

[15] R Development Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org/`.

[16] Sim, C.H. (1986). Simulation of Weibull and gamma autoregressive stationary process. *Comm. Stat.–Simul. Comput.*, **15**, 1141–1146.

[17] Tavares, L.V. (1980). An exponential Markovian stationary process. *Journal of Applied Probability*, **17**, 1117–1120.

[18] Van Wagner, C.E. (1987). Development and Structure of the Canadian Forest Fire Weather Index System. Technical Report 35, Can. For. Serv., Ottawa, Ontario.

[19] Venables, W. N. and Ripley, B.D. (2002). *Modern Applied Statistics with S.* Springer-Verlag, New York. USA.

Chapter 4

RANDOMIZING THE PROMETHEUS FIRE GROWTH MODEL

4.1 Introduction

When a wildfire is reported, fire management agencies would like to predict its future growth in order to make an appropriate resource allocation and to assure fire crew safety. In Canada, these kinds of decisions are currently based on the Canadian Fire Behaviour Prediction (FBP) system (Forestry Canada, 1992). Some agencies, such as Alberta Sustainable Resource Development, regularly use the Prometheus Wildland Fire Growth (Tymstra et al, 2010) simulator to make predictions. The Prometheus simulator is based on the FBP system combined with a mathematical model for fire spread. Inputs are weather forecast information, vegetation (fuel) and topography. The Prometheus program produces a map with overlaid contours outlining the expected fire perimeter at a specified sequence of times.

Prometheus is a deterministic simulator, but what it is modelling is highly unpredictable. Therefore, it is not realistic to expect it to give accurate predictions routinely. Accurate burn probability maps would be preferred, since they would outline areas that are likely to burn, and they would give a sense of the uncertainty in the prediction. The current approach to this problem lies in the use of Monte Carlo ensemble models which may not be satisfactorily conveying true prediction uncertainty. These methods also suffer from being too time-consuming to be operational in real-time. In the present work, we outline a statistical modelling strategy which we believe will take us a step closer to the goal of producing fast, accurate burn probability maps.

Since Prometheus is a predictive model for fire growth, it has been named after an entity associated with both forethought and fire. Adding uncertainty or randomness to its predictions leads us to an entity associated more with chaos and disorder: Dionysus.

The rest of this paper will proceed as follows. In the next section, we shall give a description of the Prometheus program and show how it is a predictive model based

on a set of experimental and wildfire data. We briefly describe a simplified research version which we have written for use in R. The subsequent section describes our statistical analysis of that data set. We then describe our proposed Dionysus algorithm, which actually comprises two approaches, both of which employ Prometheus; the first repeatedly runs randomly perturbed versions of Prometheus, and the second runs Prometheus a single time. A theoretical and empirical comparison of the two approaches is then carried out. The paper concludes with a summary of our contribution as well as what we see as the limitations of our methodology.

4.2 Background Information

The basic fire prediction question that this paper addresses is: for a given wildfire ignition, weather forecast, and fuel and topographic map, where is the fire likely to be burning in a specified period of time? It should be noted that this is only one of many issues that fire managers face, but this question is one that requires a rapid response.

Questions concerning more general fire risk are not addressed in this paper directly. For example, finding the probability that a particular area will burn in the next year or in the next 50 years is a very different question.

In fact, the Prometheus simulator has been used to try to answer both of the above kinds of questions. In the first case, it can be applied directly, but the answer will be deterministic. In the second case, randomly generated ignitions and realistic, but non-specific weather scenarios have been used as inputs to Prometheus in a Monte Carlo ensemble procedure to estimate relative frequencies of burning events. This ensemble approach has also been used to determine probabilities of burning, given a single fire ignition. The purpose of the present paper is to discuss a methodology which could provide more realistic probabilities of burning, given a single fire ignition. We will not address the question of overall fire risk further.

4.2.1 Prometheus

Prometheus is a single-platform implementation of a spatially explicit process model, derived from Huygens' principle of wave propagation. Initially, the simulated fire front is represented by a small circle centered at the ignition point. At subsequent

times, the fire front is defined as the envelope of ellipses which correspond to small fires emanating from ignition points located on the previously defined fire front.

The parameters of each ellipse (as illustrated in Figure 4.1) are based on rasterized estimates of the rate of spread (ROS) which are, in turn, related to local topography, fuel type (i.e vegetation), moisture, and wind speed and direction. Specifically,

$$a = \frac{(R+B)\Delta t}{2}; \quad b = F\ \Delta t; \quad c = \frac{(R-B)\Delta t}{2} \tag{4.1}$$

where R = forward ROS, B = back ROS, and F = flank ROS, and Δt is the time step size.

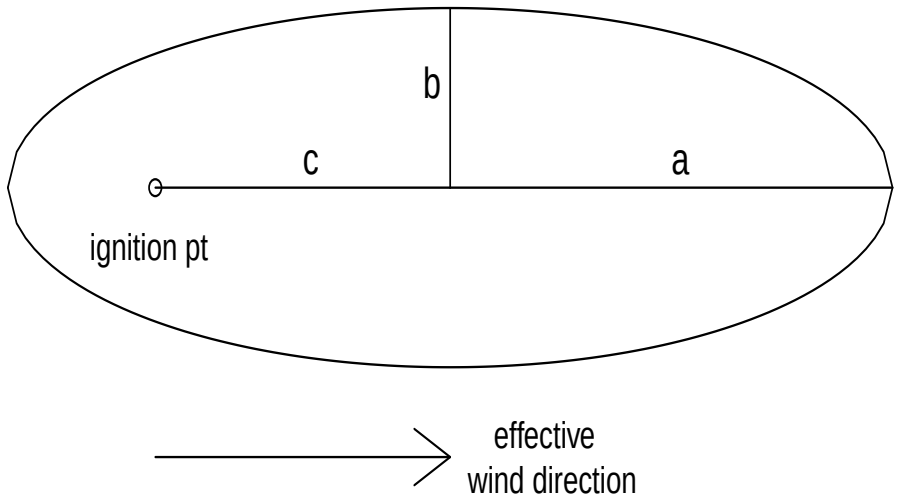

Figure 4.1: A Prometheus ellipse.

ROS values are calculated, for each grid cell, using an empirical model with interpolated forecast data, often calibrated with current observations from nearby weather stations, together with topography and fuel maps. More detail on this can be found in Tymstra et al. (2010).

The empirical ROS model describes a sigmoidal relationship between ROS and I, the *initial spread index* or ISI, a summary of the effects of both wind and moisture:

$$R = \alpha\,(1 - \exp(-\beta I))^{\gamma} \tag{4.2}$$

The α, β and γ parameters have been estimated for several different fuel types using an *ad hoc* curve-fitting procedure (Forestry Canada, 1992).

Back and flank rates of spread (BROS and FROS, respectively) are derived quantities:

$$B = \alpha \left(1 - \exp(-\beta k(w) I)\right)^{\gamma} \tag{4.3}$$

and

$$F = \frac{1}{2\rho(w)}(R + B). \tag{4.4}$$

where $k(w)$ is exponentially decreasing function of windspeed with $k(0) = 1$:

$$k(w) = \begin{cases} \mathrm{e}^{-0.10078w}, & \text{for } w < 40 \\ \frac{\mathrm{e}^{-.05039w}}{12(1-\mathrm{e}^{-0.0818(w-28)})}, & \text{for } w \geq 40 \end{cases}$$

The α, β and γ parameter values used in (4.3) are the same as those used in (4.2). The length-to-breadth ratio employed in equation (4.4) is calculated from

$$\rho(w) = 1 + 8.729(1 - \exp(-0.3w))^{2.155}$$

where w is the local effective windspeed.

Effective wind speed and direction are determined from the forecast wind speed and direction, combined with a slope effect. Details are described in the Fire Behaviour Prediction document published by Forestry Canada (1992).

The rate of spread calculations are carried out at each time step, and are performed internally to the Prometheus program, using user-supplied weather inputs (fuel moisture, wind speed, wind direction) and fuel type and topography. These (local) rates of spread are then substituted into the equations (4.1) giving parameters for all of the ellipses emanating from the most recently generated fire front.

The successive fire fronts as represented by the ellipse envelopes are governed by a system of differential equations (Richards, 1990). The Prometheus simulator program implements a numerical solution of these differential equations, calculating estimates of the envelopes at a sequence of time steps and drawing them as confocal contours on a map. A more detailed description of the Prometheus simulator, accessible to statisticians, can be found in the papers by Garcia et al (2008) and Barber et al (2008).

What should now be evident is that the Prometheus program is actually an elaborate statistical predictive model based on an underlying set of data: the predictor variable is the initial spread index, and the response variable is rate of spread. The successive ellipse envelopes that are produced as output from the Prometheus program can then be viewed as a transformation, or collection of transformations, of the response variable.

4.2.2 RPrometheus

We have programmed a slightly simplified research-version of Prometheus in R (R Core Team, 2013) which maps the burned area at each given time step. This version does not have all of the functionality of the professional version, but, because it is open-source, it is more well-suited for basic research purposes of the type we wish to pursue. In particular, the user of RPrometheus has full control of the FBP inputs, unlike the professional version. RPrometheus can be executed on a wider range of platforms. It is available at `http://www.stats.uwo.ca/faculty/braun/Rpackages.php`.

Like Prometheus, RPrometheus is based on Huygens' principle, but the implementation is somewhat different. Instead of successively, numerically, solving the differential equation system to obtain the sequence of ellipse envelopes, RPrometheus is fully rasterized. At the first time step, ellipse parameters are calculated based on fuel and weather information at the ignition site. All fuel grid cells lying within this ellipse are assumed to be ignited. At subsequent time steps, ellipses originating at previously ignited grid cell centers are obtained, and any fuel grid cells lying within these ellipses are assumed to then ignite. Note that the initial ignition might occupy more than one grid cell; in that case, the simulation begins with multiple ellipses emanating from each ignited grid cell center.

One reason for adopting a direct strategy in RPrometheus is that it avoids the issue of tangles that can occur in the ellipse envelopes. This issue is documented in Barber et al (2008) and is a side effect of the numerical method being applied to the differential equations. Coping methods have been developed (see Bose et al, 2009), but these are time-consuming and in an interpreted language like R will make simulation prohibitively slow. Because the two implementations can be made to approximate each other arbitrarily closely by appropriate selection of the time step size and raster size, we will make no further reference to RPrometheus, viewing the two implementations as essentially equivalent.

Example output from a Prometheus simulation run is displayed in Figure 4.2 starting at the ignition point of a real fire that burned in the Muskoka Lakes District of Ontario, Canada in 1999. The wind speed and fuel moisture used in the simulation correspond to the actual fire record. The wind direction was not recorded, but is taken from the west (the prevailing direction) in the simulation. The yellow portion of each panel highlights the predicted burned area at the indicated time from ignition. The blue areas of the map denote non-fuel areas; most of these areas represent water, but some may be roads or other substrates that will not burn. The green areas of the map denote regions of fuel that has not burned. In fact, the fuel grid contains a large number of types of fuel, but for simplicity of presentation, we have restricted to two colors, one for coniferous fuels (faster burning) and one (lighter green) for deciduous fuels.

The actual duration of the fire is not precisely known; the record indicates that the time between the initial report and when the fire was completely extinguished is no more than 7 days, but the number of hours of intense burning is probably much less. The simulation shows what is expected to be burned in 4 hours: an area of 61.1 hectares. The actual final fire area was 75 hectares.

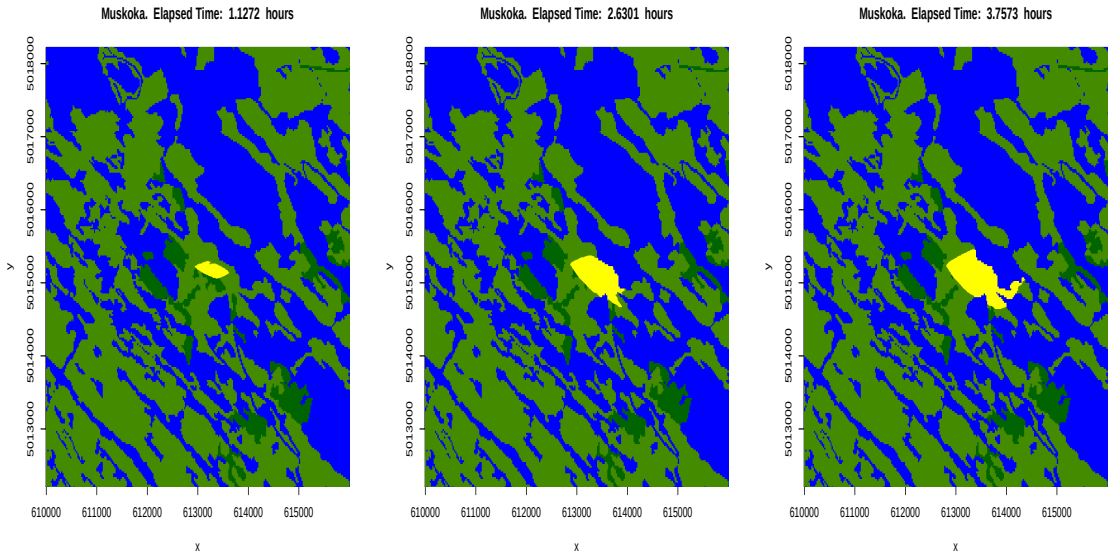

Figure 4.2: A Prometheus run in a 4km × 4km region of the Muskoka Lakes District. The yellow region represents the area burned by the simulated fire by 1.12 h, 2.63 h and 3.76 h, respectively. The horizontal coordinates are in meters easting and the vertical coordinates are in meters northing.

4.2.3 Variability and the Use of Ensemble Methods

Since Prometheus is deterministic, every run under identical conditions will produce identical output. However, uncertainty in the fire process leads to unpredictable outcomes. Cruz and Alexander (2013b) have pointed out that prediction errors in ROS can be extremely large. Thus, a realistic fire growth simulator, based on ROS, should have a stochastic component. Modelling this stochastic component is not a straightforward exercise, however.

Efforts have been made to incorporate randomness in models for ROS and for fire spread going back to Kourtz (1972). Burn-P3 (Parisien et al, 2005) represents one way in which randomness has been incorporated into Prometheus. This is an ensemble procedure based on random weather streams. Weather data are randomly selected from the historic record and used as input to the Prometheus program, resulting in a different set of simulated fire fronts each time. Empirical burn probabilities for each location are calculated by counting the number of simulations in which that location was burned divided by the total number of simulations.

Ensemble methods are also employed by Anderson et al (2007), Cruz and Alexander (2009), Finney et al (2010) and Cruz (2010) to obtain empirical burn probabilities or probability distributions for rates of spread.

4.2.4 A Limitation of the ROS Model Equations

Cruz and Alexander (2013b) have carefully laid out the principal sources of variation in the rate of fire spread which lead to prediction errors in ROS models. Clearly, input weather variables such as wind speed and direction, relative humidity, temperature and precipitation play an enormous role. Other factors that must also be acknowledged are unobservable variations in surface winds, horizontal and vertical fuel variability, the chaotic nature of turbulent flows, and feedback mechanisms associated with the fire and the atmosphere and boundary layer.

Written in its current form, the model expressed by (4.2) is limited in its ability to include stochastic variation. Variation in ROS appears to only depend on variation in the weather inputs via the initial spread index, I, hence the focus on ensemble models which randomize weather inputs. This approach will succeed only if there is no model error. Cruz and Alexander (2013a, 2013b) demonstrate clearly that ROS models, including equation (4.2), have error.

The issue may be easier to appreciate when expressed mathematically. When model (4.2) is viewed deterministically, it takes the form

$$R = g(I) \tag{4.5}$$

for the given function $g(.)$. The only way randomness can enter this model is through I, the initial spread index.

Let $\mu_I = E[R|I]$, the expected rate of spread for a given initial spread index, and let $\mu = E[R]$, the unconditional mean rate of spread. The main point of the current paper is that even if the initial spread index is known or well-approximated, there will still be substantial uncertainty in predictions made about fire spread. Thus, our interest will center on μ_I, or a conditional expectation, given the forecast error distribution. Estimating the unconditional mean may be of interest in a more general risk assessment problem, but it is not of use when predicting the behaviour of an individual fire.

Under model (4.5), the conditional variance of R is clearly

$$V(R|I) = V(g(I)|I) = 0.$$

Thus, if there is no forecast error in I, there can be no variation in R. Randomizing I as is done in the current implementation of ensemble methods will possibly address the issue of forecast error, but it cannot account for the other sources of variation described by Cruz and Alexander (2013a, 2013b).

We believe a more effective approach is to include a random term representing the uncertainty associated with the model itself:

$$R = h(I, \varepsilon) \tag{4.6}$$

where h is some as yet to be determined function, and ε is a random variable that summarizes unexplained variation. Under model (4.6), the conditional variance of R, given the initial spread index, is

$$V(R|I) = \int (h(I, x) - \mu_I)^2 f_\varepsilon(x) dx > 0$$

where $f_\varepsilon(x)$ denotes the distribution of the model error. This motivates the approach we outline in the next section: identifying a useful form for $h(I, \varepsilon)$.

4.3 A Statistical Analysis of the ROS Data

The empirical relation described by (4.2) can be augmented by a noise component which models the uncertainty in ROS. By revisiting the original data on which this relation was built, we can study this noise component. A scatter plot of the residuals from (4.2) for one of the conifer fuel types appears in Figure 4.3. Surface fires and crown fires are distinguished by different plotting characters.

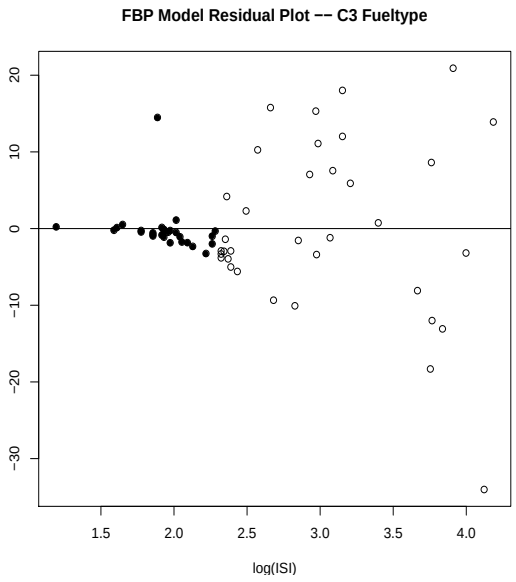

Figure 4.3: Residuals, observed rates of spread minus predicted rates of spread, using model (4.2) applied to data from historic fire records. The type of plotting character indicates fire type: black dots denote surface fires, and open circles denote crown fires.

The plot clearly indicates that the noise variance is not constant; the spread of the residuals increases with the initial spread index. The plot also indicates that ROS variability may be quite different for surface fires than for crown fires. Similar kinds of patterns can be seen for other fuel types. The nonconstancy of the variance has a serious implication; the least-squares estimates of the model parameters are inefficient, and could be improved upon with weighted least-squares or with a variance-stabilizing transformation.

We seek a variance-stabilizing power transformation:

$$R^{\delta} = (\alpha\,(1 - \exp(-\beta I))^{\gamma})^{\delta} + \varepsilon = \alpha^{\delta}(1 - \exp(-\beta I))^{\gamma\delta} + \varepsilon \tag{4.7}$$

where ε is approximately normal with mean 0 and variance σ^2. Ideally, such a transformation would be determined for each fuel type and fire type combination. However, data is sparse for many of the fuel types, so we have found it necessary to pool the data into categories which make physical sense: conifer, deciduous and mixed.

The bulk of the data are in the conifer category, so we have some confidence in the results we are about to demonstrate. For the deciduous category, there are no crown fire observations, and there are very few observations in the mixed category. In view of this, and noting that faster burning (and hence more dangerous) fires are in the conifer category, we focus on this category for the remainder of this paper.

In order to determine the transformation used in (4.7), a Box-Cox procedure (e.g. Venables and Ripley, 2002) can be used. For crown fires, the procedure suggests $\delta = 0.6$, while for surface fires, $\delta = 0$ is recommended, which translates to a log transformation on ROS. Figure 4.4 contains scale-location plots both before and after transformation which indicate that the Box-Cox recommendations should work quite well. To obtain the scale-location plots, residuals were calculated, for each conifer crown fire, using

$$e = R^{\delta} - \alpha^{\delta}\,(1 - \exp(-\beta I))^{\gamma\delta} \tag{4.8}$$

with $\delta = 0.6$, and using

$$e = \log(R) - (\log(\alpha) + \gamma \log(1 - \exp(-\beta I))) \tag{4.9}$$

for the surface fires. A plot of the square root of the absolute values of these residuals versus the corresponding predicted ROS (on the transformed scale) should show no pattern if δ is an appropriate power to use. Overlaying the scatter plot with a resistant nonparametric smooth helps to determine if there is a pattern or not. In our work, we are using Tukey's (1977) running median smoother, together with the plotting routines in the *lattice* package (Sarkar, 2008).

The leftmost panel in Figure 4.4 shows the residuals for the crown fire ROS observations when $\delta = 1$ (i.e. the untransformed data). The overlaid smooth curve shows a clear increase (consistent with what we saw in Figure 4.3).

The first panel in the right plot of Figure 4.4 shows what happens to the residuals for the coniferous crown fires when $\delta = 0.6$. The pattern disappears and the overlaid smooth curve is now essentially flat.

The second panel in the left plot of Figure 4.4 shows the residuals for untransformed surface fire ROS observations. As in the crown fire case, there is a clear need for transformation. The second panel in the right plot of Figure 4.4 shows the residuals after the log transformation is applied. Although still not perfectly constant, heterogeneity has been greatly reduced.

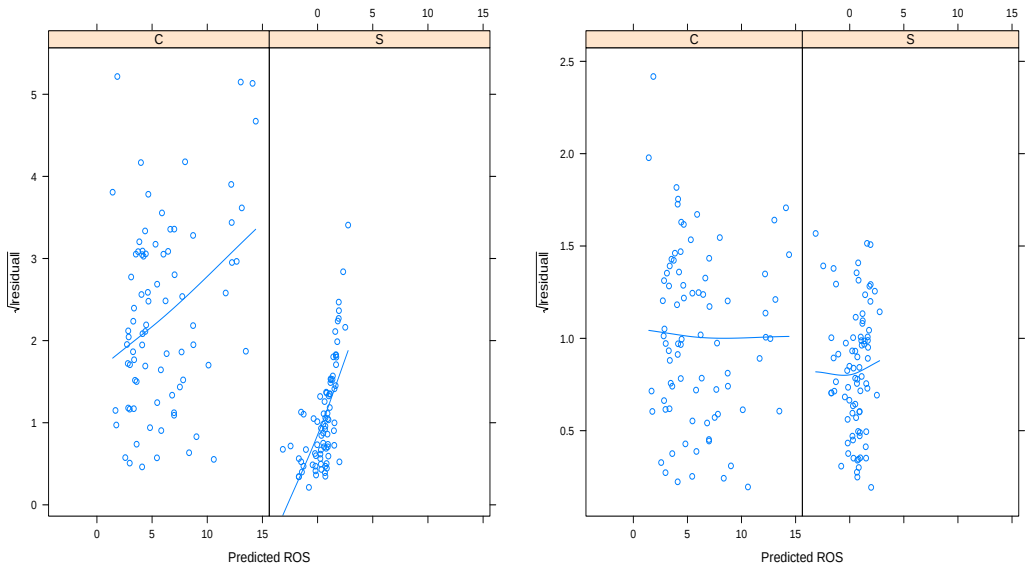

Figure 4.4: Left Plot: Residual scale-location plots for raw ROS data, with overlaid nonparametric smooth curves for crown fires (C, left panel) and surface fires (S, right panel). Right Plot: Residual scale-location plots and nonparametric curves for the same data, after transformation.

Normal QQ-plots are also of use. The upper panels in Figure 4.5 show normal QQ-plots for the residuals from the model (4.2) using untransformed surface and crown fire ROS. In both cases, the plotted points deviate substantially from the reference line, indicating non-normality. The crown fire data exhibit noticeable right-skewness, while the surface fire data are skewed heavily to the left.

The bottom row of plots shows what happens when the $\delta = .6$ power transformation and log transformation are applied to the crown fire and surface fire ROS values in formulas (4.8) and (4.9). The pattern of points on both QQ-plots is somewhat closer to linear, indicating that the transformations are normalizing as well as variance-stabilizing, especially in the case of the surface fires.

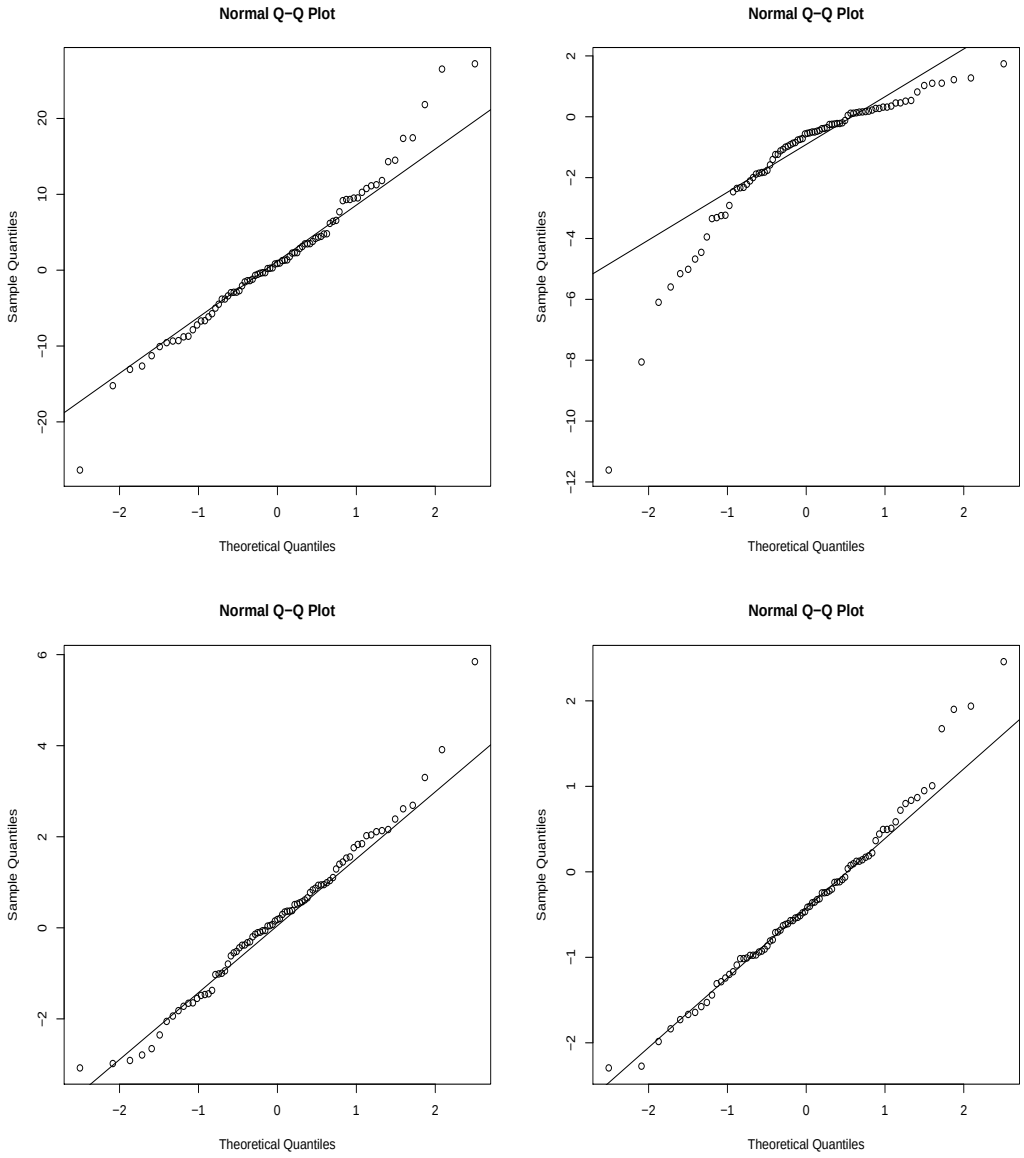

Figure 4.5: Residual normal QQ-plots for raw (top row) and transformed (bottom row) ROS data. The plots on the left correspond to crown fires; surface fires are on the right.

Based on these observations, we can provide estimates of the probability distribution of R, under the assumption of an accurate weather forecast. Specifically, the true rate of spread will be less than R_p with probability p when

$$R_p = h_\delta(I, \varepsilon_p \hat{\sigma}),$$

where

$$h_\delta(I, Z) = (\alpha^\delta(1 - \exp(-\beta I))^{\gamma\delta} + Z)^{1/\delta},$$

for $\delta \neq 0$ and

$$h_0(I, Z) = \exp(Z)\alpha(1 - \exp(-\beta I))^\gamma$$

and where ε_p is the pth quantile of the standard normal distribution. For surface fires, the error standard deviation is estimated as $\widehat{\sigma} = 0.923$ and for crown fires, $\widehat{\sigma} = 1.637$.

Variability in BROS is not expected to play a major role in the variability in the overall growth of a fire when the windspeed is high, but when windspeeds are low to moderate, BROS will be more important. The relationship at (4.3) relates the scale in the forward initial spread to the back initial spread; the formula below employs the same scaling relation to the errors in the back rate of spread. The pth percentile of BROS is modelled as

$$B_p = h_\delta(k(w)I, k(w)\widehat{\sigma}\varepsilon_p).$$

The virtue of this formula is that when the windspeed is 0, variability in BROS will match variability in ROS and FROS; this would be expected in such a case of radial symmetry. As the windspeed increases, BROS and its variation will decrease.

4.4 Dionysus: Randomized Prometheus

This section contains the Dionysus algorithm, the method for creating burn probability envelopes for a single fire modelled by Prometheus. The method will predict a shape or contour which should contain the "true" fire with probability p. This means that any site outside of the pth probability contour should burn with probability strictly less than $1 - p$, and each location inside the pth contour should burn with probability at least $1 - p$.

The ignition point (or points) of the fire is assumed known. It is also assumed that the fire type is known; that is, the fire is known to be either a surface or a crown fire, although relaxing this assumption poses little difficulty. Hourly weather is assumed known; let I_t denote the ISI at time t, and let w_t denote the corresponding windspeed.

We present two methods for producing Dionysus burn probability envelopes, an ensemble method and a percentile method. The ensemble method will require

a large amount of simulation time but will be applicable very generally. The percentile method is very fast, but it will be applicable only when the weather forecast is trustworthy.

4.4.1 Ensemble Method

Prometheus is run N times from the same ignition point. For each simulation, the time steps are $t_1, t_2, \ldots, t_m$.

For $k = 1, 2, \ldots, N$,

1. Designate one or more grid cell sites as ignited.
2. Generate ε_k^* randomly from the standard normal distribution.
3. For $t = t_1, t_2, \ldots, t_m$, do the following:
   (a) For all grid cell sites (i, j) that are ignited, do the following:
      i. Set
      $$R_{ijkt} = h_\delta(I_t, \varepsilon_k^* \hat{\sigma})$$
      where the parameters of h: α, β, γ and δ depend on the fuel type, topography and fire type at grid cell (i, j).
      ii. In the same way, set
      $$B_{ijkt} = h_\delta(k(w_t)I_t, k(w_t)\varepsilon_k^* \hat{\sigma})$$
      iii. Set
      $$F_{ijkt} = \frac{B_{ijkt} + R_{ijkt}}{2\rho(w_t)}$$
      iv. Set
      $$a_{ijkt} = \frac{(B_{ijkt} + R_{ijkt})\Delta t}{2},$$
      $$b_{ijkt} = F_{ijkt}\Delta t,$$
      and
      $$c_{ijkt} = \frac{(R_{ijkt} - B_{ijkt})\Delta t}{2}.$$

(b) Designate all sites within the union of the ellipses with parameters a_{ijkt}, b_{ijkt} and c_{ijkt} as ignited.

4. Let $\widehat{P}_{ij}$ be the proportion of simulation runs where grid site (i, j) is ignited by time t_m. The p burn probability envelope is the union of all grid sites where $\widehat{P}_{ij} \geq 1 - p$.

4.4.2 Percentile-Based Method

In this case, the Prometheus simulator is run once, starting from the given ignition. The value of p is given. For example, if a 95% burn probability envelope is desired, then p is taken to be 0.95.

1. Designate one or more grid cell sites as ignited.

2. Define ε_p as the upper pth quantile of the standard normal distribution.

3. For $t = t_1, t_2, \ldots, t_m$, do the following:

   (a) For all grid cell sites (i, j) that are ignited, do the following:

      i. Set
      $$R_{ijpt} = h_\delta(I_t, \varepsilon_p \hat{\sigma})$$
      where the parameters of h: α, β, γ and δ depend on the fuel type, topography and fire type at grid cell (i, j).

      ii. In the same way, set
      $$B_{ijpt} = h_\delta(k(w_t)I_t, k(w_t)\varepsilon_p \hat{\sigma})$$

      iii. Set
      $$F_{ijpt} = \frac{B_{ijpt} + R_{ijpt}}{2\rho(w_t)}$$

      iv. Set
      $$a_{ijpt} = \frac{(B_{ijpt} + R_{ijpt})\Delta t}{2},$$
      $$b_{ijpt} = F_{ijpt}\Delta t,$$

and

$$c_{ijpt} = \frac{(R_{ijpt} - B_{ijpt})\Delta t}{2}.$$

(b) Designate all sites within the union of the ellipses with parameters a_{ijpt}, b_{ijpt} and c_{ijpt} as ignited.

4. The p burn probability envelope is the set of all grid sites that have been designated as ignited.

4.5 Comparison of the Two Dionysus Algorithms

In this section, we compare the two Dionysus methods and demonstrate both theoretically and numerically that they give the same results when the ensemble method is run a large number of times (so as to minimize simulation error).

4.5.1 Theoretical Comparison

Let $E_n(\varepsilon)$ denote the envelope of Prometheus ellipses at time t_n generated according to one of the iterates of the algorithm described in Section 4.4.1, using the standard normal variate ε, for given forecast weather and a given fuel grid, for $n = 1, 2, \ldots, m$. Ideally, the ensemble method would be run an infinite number of times in order to reduce simulation error to 0. The pth probability contour for this version of the ensemble method is constructed from the distribution of the $E_n(\varepsilon)$'s, identifying sites which burn with probability at least $1-p$. It is this theoretical version of the ensemble method that we will compare with the percentile method.

Observe first that $E_n(\varepsilon_p)$ coincides with the pth probability envelope constructed from the percentile method. Our objective is to show that $E_n(\varepsilon_p)$ contains the pth probability contour produced by the ensemble method, with probability p. In other words, we claim that

$$P\left(E_n(\varepsilon) \subseteq E_n(\varepsilon_p)\right) = P(\varepsilon \leq \varepsilon_p) = p.$$

The following proposition provides the theoretical justification for the first equality in the above statement, since it may not be obvious that the E_n mapping possesses the needed monotonicity property.

Proposition. *Assume $E_n(\varepsilon_p)$ is the pth probability envelope produced by the algorithm of Section 4.4.2. If $\varepsilon_{p_1} < \varepsilon_{p_2}$, then $E_n(\varepsilon_{p_1}) \subset E_n(\varepsilon_{p_2})$ for all $n = 1, 2, \ldots$.*

We prove this by induction on the time step.

Suppose $\varepsilon_{p_1} < \varepsilon_{p_2}$. Then $E_1(\varepsilon_{p_1}) \subset E_1(\varepsilon_{p_2})$, since the two ellipse envelopes are just ellipses that emanate from the same ignition point, but where the ellipse parameters based on ε_{p_2} are larger than those based on ε_{p_1}.

Now, suppose $E_{n-1}(\varepsilon_{p_1}) \subset E_{n-1}(\varepsilon_{p_2})$ for some $n \geq 1$. Then, all ignition points for the ε_{p_1} ellipses must lie in $E_{n-1}(\varepsilon_{p_2})$. Therefore the set of ignition points for the set of ε_{p_1} ellipses is a subset of the set of ε_{p_2} ellipse ignition points. The parameters for the ε_{p_2} ellipses originating at ignition points for the ε_{p_1} ellipses are larger than for the ε_{p_1} ellipses. Thus, the union of the ε_{p_2} ellipses must contain the union of the ε_{p_1} ellipses, but this implies that

$$E_n(\varepsilon_{p_1}) \subset E_n(\varepsilon_{p_2}).$$

4.5.2 A Simulation Comparison

Using realistic fuel grids, burn probability envelopes calculated from the simulation-based method will be compared with percentile-based envelopes. Under homogeneous fuel, topography and weather conditions, the true burned area of a fire at a given time t will be inside the corresponding simulated area with probability p when the simulation is run with rate of spread R_p.

We use three realistic scenarios from the Muskoka Lake District of Ontario to demonstrate numerically that the percentile method produces the same output as the ensemble method. The locations were chosen randomly from among sites where ignitions have occurred in the past. Historic weather information corresponding to the given fire was used in the simulation. The wind direction was not recorded, so a west wind was assumed.

For the ensemble method examples, 500 fires were simulated at each of three locations. In each case, the proportion of times each site burned was calculated, and maps showing where the proportions exceeded 0.9,0.75, 0.5, 0.25 and 0.1 were produced. The results of the three simulations are shown in the top panels of Figures 4.6 - 4.8.

For the percentile method, simulations were conducted for the same length of time, using the same weather conditions and fuel infomations as for the Ensemble

method, but using the 0.1, 0.25, 0.5, 0.75 and 0.9 percentiles of the standard normal distribution in the calculation of ROS and BROS. The results are shown in the bottom panels of Figures 4.6 to 4.8.

In all cases, it can be seen that the ensemble method and the percentile produce very similar results as predicted by the theory. The slight differences that are evident are due to simulation error, and would be eliminated if the ensemble method could be run infinitely often. Note that the usual Prometheus output corresponds to the 50th percentile of Dionysus.

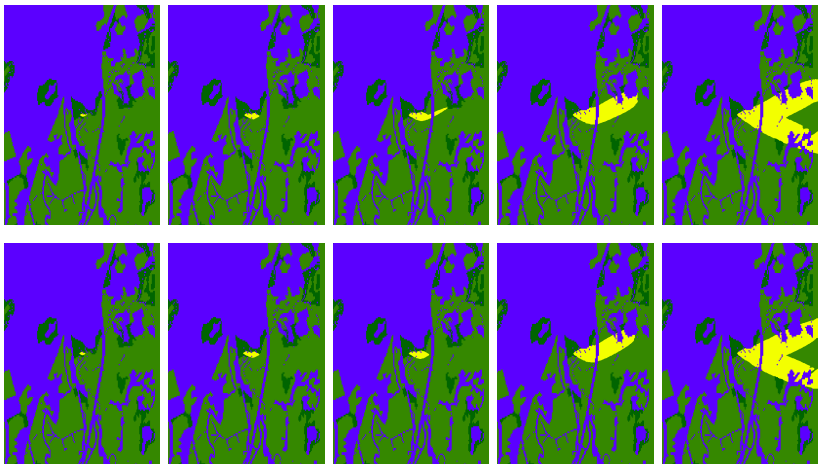

Figure 4.6: Dionysus runs in a 4 km $\times$ 4 km region in the Muskoka Lakes District. The yellow region represents the area burned by the simulated fire after 1 hour of burning. Upper panels: 10, 25, 50, 75 and 90 percent burn probability regions obtained by the simulation approach; lower panels: corresponding burn probability areas obtained by the percentile approach.

4.6 Discussion

We have proposed a statistical method for modelling fire growth prediction uncertainty based on residual variation in the ROS model. We have shown that the data originally used to develop fire behaviour prediction models also contain information about rate of spread variability which is not due to measurement or forecast error in the weather. Other forest-specific factors are causing substantial variation which

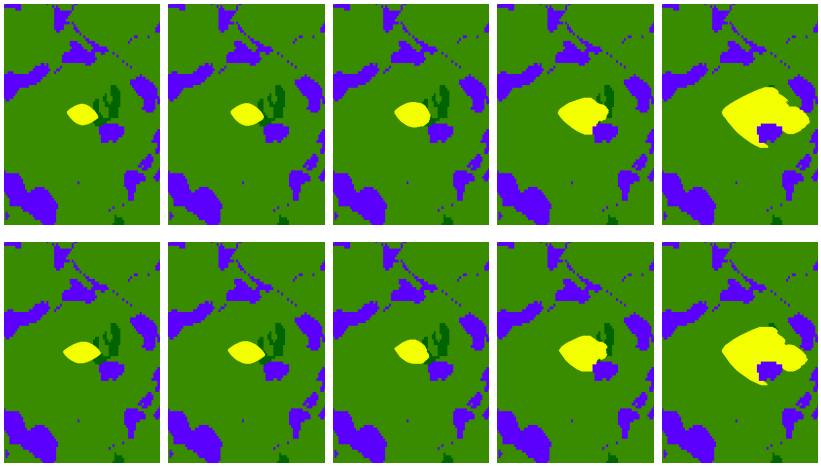

Figure 4.7: Dionysus runs in a 2 km × 2 km region in the Muskoka Lakes District. The yellow region represents the area burned by the simulated fire after 4 hours of burning. Upper panels: 10, 25, 50, 75 and 90 percent burn probability regions obtained by the simulation approach; lower panels: corresponding burn probability areas obtained by the percentile approach.

has not been taken into account in past attempts to model variation in fire growth. Our proposed Dionysus algorithm accounts for this form of uncertainty so that it can be used to produce burn probability maps, when the weather and fuel are assumed to be accurately known. There are two methods: Monte Carlo ensemble and percentile. Both give similar results. The percentile method is very fast so it can be used operationally when the weather forecast and fuel map are trustworthy.

We believe that what we are proposing in this paper represents a useful step in the important problem of assessing fire growth uncertainty, but we are also very aware that our contribution is just a step. There are serious limitations which must be acknowledged.

First, our demonstrations of the equivalence of the percentile and ensemble Dionysus methods have been based on historic data, not forecast data. Both algorithms as outlined in the paper do not explicitly account for weather forecast or measurement error. Accounting for forecast error with the percentile method will be a challenge, unless the fuel is relatively homogeneous. Incorporating forecast error should be relatively straightforward with the ensemble method, but it will require larger amounts of simulation time. Specifically, we would replace steps 3(a)i. and

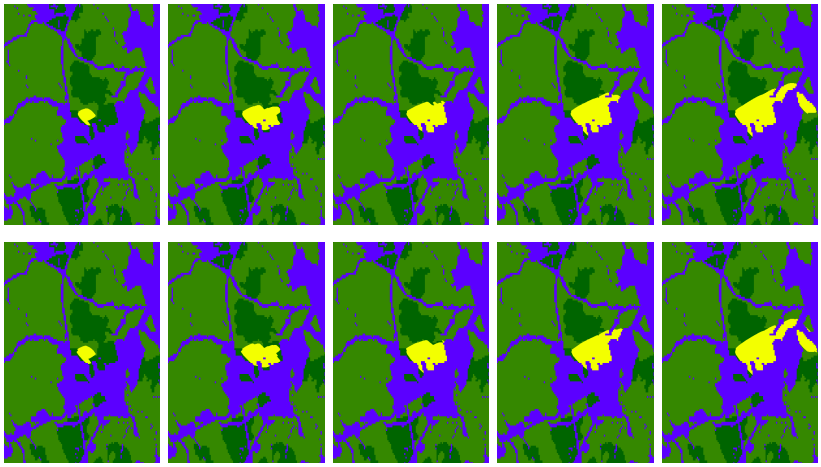

Figure 4.8: Dionysus runs in a 3 km × 3 km region in the Muskoka Lakes District. The yellow region represents the area burned by the simulated fire after 2 hours of burning. Upper panels: 10, 25, 50, 75 and 90 percent burn probability regions obtained by the simulation approach; lower panels: corresponding burn probability areas obtained by the percentile approach.

3(a)ii of the algorithm in Section 4.4.1 with the following:

i Set

$$R_{ijkt} = h_\delta(I_t + \eta^*_{kt}, \varepsilon^*_k \hat{\sigma})$$

and

ii Set

$$B_{ijpt} = h_\delta(k(w_t)(I_t + \eta^*_{kt}), k(w_t)\varepsilon_p \hat{\sigma})$$

where $\eta^*_{kt_1}, \eta^*_{kt_2}, \ldots, \eta^*_{kt_m}$ are block-resampled forecast errors (differences between actual and forecast observations) in the initial spread index. Block-bootstrapping (e.g., Davison and Hinkley, 1997) the forecast error is currently under study by the authors.

From an operational point of view, we then offer the following recommendation: the quick percentile Dionysus method should be used to make short term (i.e., less than 24 hours) predictions, since very short term weather forecasting may be sufficiently accurate. We would recommend the use of the slower ensemble form method, with resampled forecast errors, for longer term predictions.

The remainder of the discussion is concerned with our statistical approach to ROS modelling, which we believe is sound. However, its accuracy is subject to limits imposed by the underlying data which we discuss below.

The variance estimates provided here are based on a limited number of wildfire and prescribed fire observations. Within some of the fuel type categories, the sample sizes are very small. The experimental fires have been ignited under carefully controlled conditions, so there will naturally be differences in the variability observed there and what is observed in wildfires. The terrain is relatively flat in most cases; thus, predictions and our approach to prediction error depend crucially on the incorporation of the slope effect into the effective wind speed. Use of our method for error in prediction in alpine areas without prior field-testing would be inadvisable.

Extrapolation is always an issue in these kinds of situations. Predictions of ROS and its associated variability at high values of ISI will be somewhat risky without basing the model on additional wildfire data which includes measurements in this range. It should also be noted that there may be some effects due to spatial clustering in the FBP data that have not been accounted for; some of the observations in the database come from experiments conducted in neighbouring forest stands. There is bound to be a certain amount of dependence in burn behaviour bound up in such observations that one won't see in completely randomly selected fire locations. A nonlinear mixed effects model would likely lead to more accurate variance estimation; we have not pursued this here because the FBP ROS equations are firmly established in the Canadian fire management community and have provided sufficiently accurate point predictions, so our preference has been to simply assess the variability about those predictions instead of developing a new model at this point in time.

By using constrained nonlinear least-squares, we can improve upon the original *ad hoc* curve-fitting approach to model (4.2). Because of spatial clustering, we also fit a hierarchical model which improves accuracy. A full spatio-temporal model for ROS is a next step.

Re-analysis of ROS data by classifying fires as surface fires, passive crown fires and active crown fires as in Cruz and Alexander (2010) could lead to a more accurate model with more accurate variance estimates.

There is evidently measurement error in the ISI that is used in the ROS calculation. The fuel moisture value used in the calculation of the index is inferred from past weather (precipitation, temperature, relative humidity and wind), and is not measured directly at the fire site. Furthermore, the forecast weather data which would be used in practice is only available from a location which may be relatively

remote from the fire site. Rate of spread predictions may be severely biased as a result; systematic underprediction has been observed by Cruz and Alexander (2013a), and this measurement error may be at least partly responsible.

The rate of spread for a given fire is actually a moving target, because of the inherently nonstationary behaviour of an evolving wildfire (Cruz and Alexander, 2009), so the question of predicting *the* ROS for a given fire is not well-posed. Estimating typical or average ROS behaviour for a given fire would be a more realistic goal, but unfortunately, it is not clear that the underlying data provide very firm support even for this.

Application of our approach to predictions made by Dionysus needs to be field tested. Accuracy will depend on the bias in the Prometheus evolutionary model itself. The statistical method that we have proposed is more generally applicable. Any (usually nonlinear) fire growth models which have an FBP-like empirical basis will be amenable to the approach.

BIBLIOGRAPHY

[1] Barber, J., Bose, C., Bourlioux, A., Braun, J., Brunelle, E., Garcia, T., Hillen, T. and Ong, B. (2008). Burning issues with PROMETHEUS–the Canadian wildland fire growth simulation model. *Canadian Applied Mathematics Quarterly*, **16**, 337–378.

[2] Bose, C.J., Bryce, R., Dueck, G. (2009). Untangling the PROMETHEUS nightmare. In Anderssen, R.S., R.D. Braddock and L.T.H. Newham (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009. pp. 240-246. ISBN: 978-0-9758400-7-8. `http://www.mssanz.org.au/modsim09/A4/bose.pdf`

[3] Cruz, M. (2010). Monte Carlo-based ensemble method for prediction of grassland fire spread. *International Journal of Wildland Fire*, **19**, 521–530.

[4] Cruz, M. and Alexander, M. (2009). Assessing discontinuous fire behavior and uncertainty associated with the onset of crowning [abstract], Page 20 in R.E. Masters, K.E.M. Galley and D.G. Despain (eds.). The '88 Fires: Yellowstone and Beyond, Conference Proceedings, Tall Timbers Miscellaneous Publication No. 16, Tall Timbers Research Station, Tallahassee, Florida, USA.

[5] Cruz, M.G. and Alexander, M.E. (2013). Uncertainty associated with model predictions of surface and crownfire rates of spread. *Environmental Modelling & Software*, **47**, 16–28.

[6] Cruz, M.G. and Alexander, M.E. (2014). Limitations on the accuracy of model predictions of wildland fire behaviour: a state-of-the-knowledge overview. To appear in *The Forestry Chronicle*.

[7] Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press,

[8] Forestry Canada Fire Danger Group (1992). Development and structure of the Canadian Forest Fire behaviour Prediction System. Forestry Canada, Ottawa, Ontario. Information Report ST-X-3.

[9] Garcia, T., Braun, W.J., Bryce, R. and Tymstra, C. (2008). Smoothing and bootstrapping the PROMETHEUS Fire Spread Model. *Environmetrics*, **19**, 836–848.

[10] Kourtz, P. (1972). Probability makes fire danger index more reliable. *Fire Control Notes*, **33**, 11–12.

[11] Parisien, M. A., Kafka, V. G., Hirsch, K. G., Todd, J. B., Lavoie, S. G., and Maczek, P. D. (2005). Using the Burn-P3 simulation model to map wildfire susceptibility, *Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre*, Edmonton, AB. Information Report NOR-X-405.

[12] R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`.

[13] Richards G. (1990). An elliptical growth model of forest fire fronts and its numerical solution. *International Journal for Numerical Methods in Engineering* **30**, 1163–1179.

[14] Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R.* Springer, New York, USA.

[15] Tukey, J. W. (1977). *Exploratory Data Analysis.* Addison-Wesley, Reading, USA,

[16] Tymstra, C., Bryce, R.W., Wotton, B.M., Taylor, S.W., and Armitage, O.B. (2010). Development and structure of Prometheus: the Canadian Wildland Fire Growth Simulation Model – Information Report NOR-X-417, Canadian Forest Service, URL `http://www.firegrowthmodel.com/`.

[17] Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S.* Fourth edition. Springer, New York, USA.

Chapter 5

CONCLUSIONS AND FUTURE RESEARCH

Statistical methods are useful in wildfire management and prediction. In this thesis, there were two main statistical techniques used to analyze FWI data. One is the nonparametric block bootstrap and the other is nonnegative time series models, which can be used in parametric bootstraps, for example. For modeling fire growth and incorporating randomness, a parametric bootstrap is used. Thus, a unifying statistical theme in this thesis is the bootstrap.

The nonparametric block bootstrap is a good way to handle time series since it makes limited model assumptions. However, confidence intervals from the block bootstrap are not accurate without adjustment. The α-level adjustment does not work as well as the proposed length-adjustment calibration method.

We saw that with length-adjustment, the confidence intervals for the variance of the first lag autocorrelation indicated changes in this correlation from year to year at several Ontario weather stations. In order to give more precise length-adjusted confidence intervals, block size selection should be carefully studied in the future. A practical blocksize selector based on a double bootstrap was subject to a preliminary investigation and it seems reasonable. However, it needs further study.

Starting from an existing exponential tailed minification time series model, a random coefficient mixture Weibull tailed minification model was developed through a sequence of data analysis steps. The stationarity of this new process was studied in a new way, using techniques from ordinary differential equations theory.

Seasonality was added to the model in agreement with earlier research on this time series (Martell, 1999). This model could be used as the basis of a parametric bootstrap simulator. The lag one autocorrelation behaviour at some of the weather stations confirmed the result obtained with the nonparametric block bootstrap: the autocorrelations appear to change from year to year.

The plots of FWI exceedances and fire occurrences at six locations show that using model-based probabilities of FWI minima over consecutive days to set fire danger cut-offs can be more effective than using single day cut-offs. Association of FWI and fire occurrence should be investigated more thoroughly so that we can better predict potential fire danger.

Dionysus is a parametric block bootstrap for generating probability contours for given fires that is a completely different method from existing simulators. Again, data analysis played a principal role in the development of Dionysus. As in the FWI modelling exercise, residuals and their visualization were useful.

Monte Carlo ensemble and percentile methods can both be employed. The percentile method is more likely to be used operationally if the weather forecast is trustworthy because it runs fast. Further validation is required: comparisons between Dionysus output and actual fire perimeters must be undertaken.

Weather forecast error is part of the uncertainty in the whole fire growth process (Cruz and Alexander, 2013, 2014) and needs to be modelled before being incorporated into either the ensemble or percentile forms of Dionysus. In addition, work on ROS modelling using covariate information related to fuel type, instead of the traditional nominal fuel types, as described by Cruz and Alexander (2010), will allow Prometheus and Dionysus to be used more generally in Canada and around the world.

BIBLIOGRAPHY

[1] Cruz, M.G. and Alexander, M.E. (2010). Assessing crown fire potential in coniferous forests of western North America: a critique of current approaches and recent simulation studies. *International Journal of Wildland Fire*, **19**, 377–398.

[2] Cruz, M.G. and Alexander, M.E. (2013). Uncertainty associated with model predictions of surface and crownfire rates of spread. *Environmental Modelling & Software*, **47**, 16–28.

[3] Cruz, M.G. and Alexander, M.E. (2014). Limitations on the accuracy of model predictions of wildland fire behaviour: a state-of-the-knowledge overview. To appear in *The Forestry Chronicle*.

Appendix A

LIST OF ABBREVIATIONS

AR	Autoregressive
BROS	Back Rate of Spread
BUI	Buildup Index
DC	Drought Code
DMC	Duff Moisture Code
EAR	Exponential Autoregressive
EMA	Exponential Moving Average
ET	Exponential Tailed
WT	Weibull Tailed
FBP	Fire Behaviour Prediction
FFMC	Fine Fuel Moisture Code
FROS	Flank Rate of Spread
FWI	Fire Weather Index
i.i.d.	independent and identically distributed
ISI	Initial Spread Index
NEAR	Newer Exponential Autoregressive
ROS	Rate of Spread
w.p.	with probability

VITA

Name

Lengyi Han

Education

DOCTOR OF PHILOSOPHY (Candidate)
in Statistics. The University of Western Ontario, London, Ontario. Supervisor: Dr. John Braun, 2009 to present. Expected completion in early 2013

MASTER OF SCIENCE
in Statistics. University of Western Ontario, London, Ontario. Supervisor: Dr. D. Bellhouse. August, 2006 to August, 2007.

MASTER OF ARTS
in Education. Shenyang Normal University Liaoning, China. Supervisor: Dr. Yunxiang Liu. September, 1995 to July, 1998.

DIPLOMA
in Computer Science. University of Science and Technology Liaoning Liaoning, China. September, 1988 to July, 1991.

Publications in Refereed Journals

1. **Han, L.S.** and Braun, W.J. (2014). Block Bootstrap Calibration with Application to the Fire Weather Index. Accepted by *Communications in Statistics – Simulation and Computation*.

2. **Han, L.S.** and Braun, W.J. (2013). Dionysus: A Stochastic Fire Growth Scenario Generator. *Environmetrics*. DOI: 10.1002/env.2236

Scholarships and Academic Awards

Scholarships

1. MITACS, The University of Western Ontario. (2009). $15,000.

2. MITACS, The University of Western Ontario. (2013). $15,000.

3. Western Graduate Research Scholarships.
   The University of Western Ontario (2009-2013).

Awards

1. Special Support Award. Faculty of Science, University of Western Ontario. 2009 - 2012. ($7000/year.)

2. Graduate Thesis Research Award. Faculty of Science, University of Western Ontario, 2012. ($1500.)